High Quality Graphene for Magnetic Sensing

Von der Fakultät Mathematik und Physik
der Universität Stuttgart zur Erlangung der Würde eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

vorgelegt von
Sven Patrick Herlinger
aus Stuttgart

Hauptberichter:	Dr. Jurgen Smet
Mitberichter:	Prof. Dr. Peter Michler
Vorsitzender:	Prof. Dr. Hans Peter Büchler
Tag der mündlichen Prüfung:	09.05.2022

Max-Planck-Institut für Festkörperforschung
Stuttgart, 2022

For my beloved mother
Thank you for everything

Table of contents

Chapter 1

Introduction

The research of graphene, a single atomic layer of graphite, has evolved at an unprecedented speed. Although this two-dimensional system played an important role for several decades as theoretical model to describe other carbon allotropes, for instance carbon nanotubes, it took until 2004 for graphene sheets to be isolated successfully and investigated experimentally. In particular the straightforward method of mechanically exfoliating such flakes from bulk graphite by means of a sticky tape helped to literally initiate a gold rush in the condensed matter physics community. This culminated in the Nobel Prize in Physics awarded to Andre Geim and Konstantin Novoselov in 2010, six years after their first publication on isolated graphene. Due to graphene's peculiar, linear band structure a multitude of novel physical phenomena, such as Klein tunneling or electron hydrodynamics, could be studied experimentally for the first time. Moreover, graphene's mechanical, optical and electrical properties are remarkable and frequently can be described only by superlatives. It possesses the highest ever measured in-plane tensile strength while showing high flexibility out of plane. It is optically transparent over a wide spectral range. It can be considered as a zero band gap semiconductor and exhibits record charge carrier mobility values at room temperature. The charge carrier density in graphene can be easily tuned over a wide density range by means of electrostatic gating while it remains relatively well conducting even at very low charge carrier densities. Around the so-called charge neutrality point (CNP), when the average charge carrier density approaches zero, the system breaks up into a spatially inhomogeneous landscape of electron and hole puddles. At higher temperatures, thermally activated charge carriers may play a role, too. Accordingly, graphene constitutes a beautiful textbook example for a system that can be tuned from a single carrier type regime at high density to a two carrier type system in the vicinity of the CNP. The mechanical exfoliation technique was also applied to similar material systems like graphite that consist of many single layers held together by van der Waals forces. As a result, a

multitude of other two-dimensional (2D) materials could be fabricated which extended the research field significantly. Eventually, the improvement of the handling process for these 2D materials allowed stacking and combining them at will in order to assemble artificial van der Waals heterostructures with engineered physical properties.

The exponential surge of scientific publications was accompanied by a similar steep rise in patent applications. Research and development departments of larger corporations soon became interested and started to explore potential applications of graphene (and other 2D materials) in electronics, opto-electronics, sensors and composite materials. Various governmental programs worldwide fueled and still fuel this development with significant amounts of public funding. One of the largest initiatives, the Graphene Flagship program initiated in 2013 by the European Commission, even granted a budget of one billion Euro over a time period of ten years for basic research and industrialization of graphene. However, today, 18 years after its discovery and after more than one decade of intensive development, there are still no graphene-based products in the market that make really use of its remarkable physical properties and have notable commercial traction or are produced in high volume. Of course, it happens often that the industrialization of a new technology eventually takes longer than anticipated. In particular, in the semiconductor and photonics industry. However, the slow transition into actual products also originates from the fact that the large scale fabrication of *high quality* graphene devices is still an unsolved challenge. As a consequence, the full potential of graphene, its very promising physical properties, cannot yet be leveraged in today's applications. This makes it more difficult for graphene to compete with established material systems, such as for instance Si or GaAs. These incumbent technologies will only be replaced, requiring significant investments to do so, when a new technology offer superior performance or is cheaper (ideally both). Or, when this new technology allows to do something not possible at all using existing technologies. However, such a "killer" application for graphene has not yet been identified. Also, the performance of graphene limited to the rather low quality achievable by today's large scale production is typically not superior to existing technologies. Moreover, the integrability of graphene with Si CMOS demands further development.

Graphene is all surface. This means that this single atom thick layer is fully exposed to its environment. As a consequence, in particular its electrical properties are strongly affected by the substrate it is placed on or by the material it is encapsulated with. Also residues from device fabrication steps or air-borne adsorbates can have a significant impact. Therefore, it is very challenging to reliably produce high quality graphene devices (of large area). When looking at a large ensemble of samples, one typically finds a significant variation in performance. Frequently

used substrates such as SiO_2 deteriorate the electrical properties of graphene, reducing the obtainable room temperature mobility drastically. It then does not even beat GaAs. All on its own graphene cannot accomplish much more than existing materials. It only reveals its remarkable properties and disruptive potential when appropriate measures are applied such as placing graphene on suitable substrates and optimizing the device fabrication steps. Excellent device quality can be achieved when encapsulating graphene between multilayer hexagonal boron nitride (hBN). This close cousin of graphene, which can also be classified as 2D material, serves as atomically flat substrate with very low charge impurity density. Unfortunately, the large scale fabrication of hBN lacks behind that of graphene. Therefore, it is important to develop a common platform for 2D materials and then industrialize this platform. In order to justify this huge (financial) effort and technical risk, several use cases must be identified that all would benefit from such common platform for high quality 2D materials, i.e. enabling products that are superior to the established competition.

One of the promising use cases that benefits most from graphene having a high quality and superior electrical properties is magnetic sensing based on Hall elements. High quality graphene devices can be tuned to very low charge carrier densities, corresponding to huge Hall coefficient values or magnetic field sensitivity, while the system remains well conducting due to the record mobility at room temperature. This dissertation investigates the reliable fabrication of high quality graphene and its use as Hall transducer material. Since "insight must precede application", as already postulated by Max Planck, we try to understand what the critical factors are that limit the performance of graphene. We then can develop methods to avoid or at least mitigate them. Finally, it is important to gain a deep understanding of the physics of the device used for the envisioned application. In our case, this means to find a suitable model to describe graphene Hall elements. In this work, we will fabricate, characterize and study a large number of devices, comparing different substrates and fabrication methods. This allows us to substantiate such model through a large data set and provide proper statistics that help to understand all relevant dependencies on as well as between the operating parameters.

We begin with a comprehensive review of the theory and experimental studies of electronic transport in graphene in Chapter 2. The peculiar, linear band structure is introduced. This includes a treatment of many body interactions that lead to a renormalization of the Fermi velocity. We then discuss the coexistence of electrons and holes around the CNP due to thermal activation as well as disorder induced spatial density inhomogeneities. We show that the channel approximation model which introduces only one ad-hoc parameter, the residual total charge carrier density at the CNP, can be used to describe the crossover from the single to the two carrier

type regime. A review of all relevant sources of extrinsic disorder is given which is also important to understand diffusive transport in graphene. We introduce the two carrier type Drude model which serves us, in combination with the channel approximation, as one of the main building blocks to describe graphene Hall elements. For the sake of completeness, the ballistic and viscous transport regimes are briefly considered as well. In addition, a summary of the quantum effects in transport, e.g. quantum interference and quantum Hall effects, is provided in Appendix B. The latter, however, will play no role for our measurements that are performed at small magnetic fields and primarily near room temperature.

In Chapter 3, different production, handling and device processing steps are described and compared. We provide sample statistics where we distinguish between graphene devices supported by SiO_2 and hBN-encapsulated devices. Both Raman and electrical transport measurements in the absence of a magnetic field are performed to identify the main extrinsic sources of disorder and show how they can be minimized. Moreover, we introduce a reliable, post-processing technique to improve the quality of hBN-encapsulated graphene even further and increase the yield of devices with highest quality. This is supported by statistical evidence in the Raman and transport measurements. Finally, we show that we reach the intrinsic limits of graphene after applying this method. The best obtainable mobility, limited by phonon scattering, as well as the smallest possible residual total charge carrier density at the CNP, limited by thermal activation of charge carriers, are determined at room temperature.

In the following Chapter 4, we discuss magneto-transport measurements in classically weak magnetic fields, i.e. when $\mu^2 B^2 < 1$ holds. The Hall coefficient is determined from the measured Hall voltage as a function of magnetic field. We demonstrate that the dependence of the Hall coefficient on the gate voltage, i.e. on the gate controlled single carrier density, can be described by the two carrier type Drude model in combination with the channel approximation. This allows to extract the residual total charge carrier density at the CNP as fitting parameter. Again, we compare different fabrication methods and study the beneficial effect of our proposed post-processing technique to improve hBN-encapsulated graphene devices. Subsequently, the gate voltage dependence of the Hall coefficient and the resistivity is studied for several temperatures between liquid helium and room temperature. From these data, the thermal activation of charge carriers is obtained and compared to the expected dependence based on a linear band structure. Finally, we make a short excursion by repeating the temperature dependent measurements for bilayer devices and studying the Moiré effect that emerges when the crystallographic planes

of graphene and hBN are well aligned. For the sake of completeness, the magneto-resistance at the CNP and potential transport phenomena beyond the diffusive picture are discussed as well.

Finally, a comprehensive technical assessment of graphene based Hall elements is conducted in Chapter 5. At first, a broader context is given by reviewing different types of magnetic field sensing in commercial applications. We then discuss Hall sensors and their key performance indicators, such as sensitivity, magnetic field resolution or power consumption, in detail. The magnetic field resolution, i.e. the minimum detectable magnetic field, is determined by the noise voltage and the sensitivity, i.e. the Hall coefficient times the applied current. Hence, we have to determine accurately the electrical noise in graphene Hall elements. For this purpose, we introduce our measurement setup which allows us to determine both the Hall coefficient as well as the noise voltage spectral density as a function of all relevant operating parameter, i.e. density, frequency and bias current. Several measures, such as a home-built sample box with integrated amplifiers and carefully selected instruments, are applied to obtain a very low instrumental noise floor of the electrical setup in order to obtain very clean and accurate noise voltage spectra from our samples. A systematic investigation of the entire parameter space then enables us to study all dependencies on these parameters as well as the interplay between thermal and 1/f noise. We propose a model based on only four material parameters to estimate the sensitivity and magnetic field resolution and simulate their functional dependence on all operating parameters. This allows predicting the performance of any type of graphene Hall element in terms of footprint, geometry and device quality (i.e. type of substrate). Typical values for all key performance indicators are given both for graphene devices on SiO_2 as well as hBN-encapsulated graphene devices. This specification sheet of graphene Hall elements is then compared to other materials which are frequently used in Hall elements, such as Si, GaAs or InSb. At the end of the chapter, we review the status-quo of large-area fabrication of graphene and thin hBN films as ideal substrate. Also, other remaining challenges such as Si CMOS integration are discussed. Finally, we critically discuss the probability and the realistic time scale of graphene becoming industrialized for high volume applications such as Hall sensors.

Chapter 2

Electronic transport in graphene

Graphene is a single atomic layer of graphite and consists solely of sp^2-hybridized carbon atoms which are arranged in a hexagonal honeycomb lattice and host a native two-dimensional electron system. Although graphene was theoretically studied for several decades as starting point to describe graphitic systems, e.g. carbon nanotubes, its systematic experimental realization was first achieved at the University of Manchester in 2004 [1]. It soon gained tremendous interest since this first truly two-dimensional material possesses an unusual, linear band structure (at low energies). This promises a plethora of new and unusual electronic properties and effects.
The peculiar feature of the band structure is also referred to as Dirac cone, since it is described by a Hamiltonian for massless chiral Dirac fermions. The two Dirac cones corresponding to the conduction and valence bands touch each other exactly at the cone tips. Hence, graphene can be considered as a zero band gap semiconductor (or semimetal). This particular touching point is then referred to as Dirac point or charge neutrality point. The latter nomenclature reflects the fact that the electronic states in pristine graphene are filled up to the Dirac point (corresponding to exact half-filling of the band structure at $E = 0$). As a consequence, at zero temperature, no free carriers are present when the system's energy is at the Dirac point, e.g. it is undoped. However, at finite temperature, due to the symmetry of the band structure, equal amounts of electrons and holes are thermally activated which motivates the term charge neutrality point (often abbreviated as CNP).

This chapter is intended to review the key experimental results and theories focusing on electronic transport in monolayer graphene both in the absence and presence of classically weak magnetic fields. Recent reviews [2–6] and books [7–9] partially covering this topic, have been published more than five years ago. They therefore have not been able to capture the most recent progress regarding the experimental realization of graphene with extremely low extrinsic disorder by the end of 2013 [10]. Since then, important refinements in the theoretical picture of diffusive transport,

renormalization effects by electron electron interactions, and mesoscopic transport effects have been made. The latter are captured by a hydrodynamic approach with first reviews appearing while this chapter is being written [11, 12].

The first section of this chapter introduces the lattice and band structure in the tight-binding model and briefly reviews the main implications of massless chiral charge carriers. Subsequently, a short overview is given on the reconstruction of the band structure caused by weak periodic potentials and by many-body interactions. The latter become important at low density due to weak screening. In the second section, we show that the charge carrier density in graphene as a zero band gap semiconductor can be easily tuned in a controlled manner by gating and unintentionally by external disorder, which imposes a spatially inhomogeneous landscape of electron and hole puddles. Such spatial inhomogeneity is a fundamental difference to conventional bulk semiconductors, which possess a band gap. Further, the interplay of disorder and thermal generation of carriers around the charge neutrality point (CNP) will be discussed. The important third section covers the diffusive transport of charge carriers under an external electric and classically weak magnetic field. We will emphasize that graphene has to be considered as a (spatially inhomogeneous) two charge carrier type system near (average) charge neutrality and present possible theoretical models. In the vicinity of the CNP, disorder not only induces a spatially fluctuating charge carrier density, but also causes predominantly scattering of the charge carriers (for $\mathrm{T} \to 0$). The microscopic origin of the disorder is still under debate. This section provides an overview of momentum relaxing scattering mechanisms including phonon scattering at elevated temperature. The temperature dependence of transport near the CNP is non-trivial: Thermal activation of charge carriers, inelastic scattering by phonons, and elastic disorder scattering (due to temperature dependent screening), all contribute to the temperature dependence. When external disorder can be reduced significantly in manufactured devices, carrier density fluctuations decrease. Then, the intrinsic physics near the Dirac point becomes accessible. It bears insulating behavior ($\partial\rho/\partial\mathrm{T} < 0$). The fourth section studies the mesoscopic transport regime in graphene which arises for low extrinsic disorder. The latter is split into a conventional ballistic regime at low temperature and high densities and into a viscous regime at intermediate temperature and densities, where carrier-carrier collisions dominate and favor a hydrodynamic model. In the last section of this chapter we summarize the discussed transport features and compare graphene to other technologically relevant materials.

Additional quantum transport effects are discussed in the Appendix B.1. They are not explicitly the subject of this thesis, since our experiments are not performed in quantizing magnetic fields. Moreover, corrections to the Boltzmann theory of

diffusive transport by quantum interference effects are negligible even at low temperature. Nonetheless, quantum interference effects do provide valuable information about the dominating microscopic disorder mechanism.

2.1. Lattice and band structure

2.1.1. From carbon to graphene to van der Waals materials

Carbon is a light element and forms strong covalent bonds in a very versatile manner by hybridization of its atomic orbitals (2s, $2p_x$, $2p_y$, and $2p_z$). They are occupied by four valence electrons. In graphitic materials such as graphene, the in-plane orbitals $2p_x$ and $2p_y$ mix with the 2s orbital resulting in three sp^2-hybridized orbitals per atom. The hybridized orbitals allow carbon to form strong σ-bonds with three neighboring carbon atoms that align in plane and are oriented 120° with respect to each other. These strong σ-bonds award a record in plane tensile strength to graphene, while still allowing a high flexibility out of plane [13]. The remaining $2p_z$ atomic orbital points out of plane and is occupied with one electron per carbon atom. The $2p_z$ orbitals unify to form a delocalized π-bond orbital cloud above and below the hexagonal plane (in analogy to the benzene ring). These delocalized π-orbitals then host a two-dimensional quasi-free electron gas which obviously can be influenced facilely by the environment of graphene, e.g. by residues on its surface and by the substrate it resides on. When two or more layers of graphene are stacked on top of each other, remnant interaction between two vertically adjacent π-orbitals results in comparatively weak van der Waals forces holding the layers together. Many layers then form the graphite crystal. The van der Waals forces are so weak that they can be overcome even by macroscopic mechanical forces during micromechanical cleavage of a graphite crystal with a scotch tape [14]. At first, this has been exploited to fabricate graphene [1]. Subsequently, other layered crystals held together by van der Waals forces were identified and the mechanical exfoliation technique was applied successfully to isolate single or few layers of different materials. Graphene was the driving force and opened a much broader research area referred to as van der Waals materials or van der Waals heterostructures. Such heterostructures can be engineered at will by combining different layered materials in order to obtain new and exciting physical properties (see Ref. 15 and Chapter 3).

2.1.2. Structure in real and reciprocal space

The hexagonal lattice of ideal graphene in real space is illustrated in Figure 2.1a. The honeycomb lattice consists of two interwoven triangular sublattices of inequivalent carbon atoms A and B. This inequality stems from the different bond directions and the two sublattices are related to each other by inversion symmetry. Accordingly, the honeycomb structure itself with an inter-atomic distance $d_{c-c} = 0.142\,\mathrm{nm}$ is not a proper Bravais lattice. The unit cell of graphene contains two carbon atoms A and B given by the shaded rhombus in Figure 2.1a. It forms the basis of a hexagonal Bravais lattice indicated by blue crosses. The Bravais lattice vectors are given by

$$\mathbf{a_1} = \left(\frac{a}{2}\,,\,\frac{\sqrt{3}a}{2}\right) \quad \text{and} \quad \mathbf{a_2} = \left(\frac{a}{2}\,,\,-\frac{\sqrt{3}a}{2}\right)\,, \tag{2.1}$$

where $a = |\mathbf{a_1}| = |\mathbf{a_2}| = \sqrt{3}d_{c-c} = 0.246\,\mathrm{nm}$ is the lattice constant [7]. In reciprocal space, the hexagonal symmetry of the Bravais lattice is conserved and the corresponding reciprocal lattice vectors are

$$\mathbf{b_1} = \left(\frac{2\pi}{a}\,,\,\frac{2\pi}{\sqrt{3}a}\right) \quad \text{and} \quad \mathbf{b_2} = \left(\frac{2\pi}{a}\,,\,-\frac{2\pi}{\sqrt{3}a}\right)\,. \tag{2.2}$$

Therefore, the first Brillouin zone of graphene is hexagonal as depicted in Figure 2.1b. It contains four pivotal symmetry points: The zone center $\Gamma = (0\,,\,0)$, the saddle point $M = \left(\frac{2\pi}{3a}\,,\,0\right)$, and the two inequivalent corners of the Brillouin zone

$$K_\xi = \left(\frac{2\pi}{3a}\,,\,\xi\frac{2\pi}{3\sqrt{3}a}\right) \quad \text{with} \quad \xi = \pm\,. \tag{2.3}$$

The latter are of particular importance for the electronic properties of graphene. They are referred to as Dirac points for reasons that will be explained in the next paragraphs. Note that all other saddle points and Dirac points in the first Brillouin zone are given by translation of the above given coordinates by the reciprocal lattice vectors. Any other arbitrary position can be described in the reciprocal coordinates $\boldsymbol{\kappa} = (\kappa_x, \kappa_y)$ measured from the zone center.

The hexagonal structure in real and reciprocal space can be observed experimentally by transmission electron microscopy and electron diffraction measurements on partially suspended graphene layers (see Figure 2.1c). Careful electron diffraction studies have revealed that suspended graphene tends to be slightly corrugated [16]. Such intrinsic ripples can explain the thermodynamic stability of graphene by making it a non-perfect two-dimensional crystal [16, 17]. On a substrate, graphene conforms to the substrate surface roughness due to its high out-of-plane flexibility and

due to the interaction of its π-orbitals with dangling bonds of the substrate (see also Chapter 3). In both cases, suspended graphene and graphene on a substrate, the above described ideal lattice has to be revised and a modulation of the bond lengths has to be considered. The latter is accompanied by random strain fluctuations. Such fluctuations cause variations in the hopping integrals [17] and represent an additional source of disorder as discussed in Section 2.3.2 [2, 18].

2.1.3. Tight binding band structure

The (opto-)electronic properties of graphene are fully determined by the π-bands since the bonding σ and the anti-bonding σ^* bands exhibit an energy gap of about 12 eV. Hence, they can be neglected in the band diagram [19]. The π-bands can be calculated in a tight binding model which approximates the wave function of the crystal by the superposition of atomic wave functions [7, 20]. The tight binding model neglects many-body interactions and we assume the ideal graphene lattice case as described in the previous section, ignoring effects of substrate roughness and strong disorder. The unit cell of graphene contains two carbon atoms, one of each sublattice A and B. This results in 2x2 matrix elements with diagonal (AA, BB) and off-diagonal elements (AB,BA). Each carbon atom has three nearest neighbors belonging to the other sublattice and in total six next nearest neighbors of the same sublattice. The transfer integral matrix elements H and the overlap integral matrix elements S of monolayer graphene read then as follows [7]

$$\mathrm{H} = \begin{pmatrix} 0 & -t_0\, f(\boldsymbol{\kappa}) \\ -t_0\, f^*(\boldsymbol{\kappa}) & 0 \end{pmatrix}, \quad \mathrm{S} = \begin{pmatrix} 1 & t_1\, f(\boldsymbol{\kappa}) \\ t_1\, f^*(\boldsymbol{\kappa}) & 1 \end{pmatrix}. \tag{2.4}$$

Here, the nearest neighbor hopping energy $\mathrm{t_0} \approx 3\,\mathrm{eV}$ and the next nearest neighbor hopping energy $\mathrm{t_1} \approx 0.1\,\mathrm{eV}$ are introduced [21]. The energy dispersion for the conduction (+) and the valence band (−) is then given by

$$\mathrm{E_\pm} = \frac{\pm \mathrm{t_0}|\mathrm{f}(\boldsymbol{\kappa})|}{1 \mp \mathrm{t_1}|\mathrm{f}(\boldsymbol{\kappa})|} \approx \pm \mathrm{t_0}|\mathrm{f}(\boldsymbol{\kappa})|\,, \tag{2.5}$$

where $\mathrm{f}(\boldsymbol{\kappa}) = \mathrm{e}^{\mathrm{i}\kappa_\mathrm{y}\mathrm{a}/\sqrt{3}} + 2\mathrm{e}^{-\mathrm{i}\kappa_\mathrm{y}\mathrm{a}/2\sqrt{3}} \cos(\kappa_\mathrm{x}\mathrm{a}/2)$. The nearest neighbor approximation was deployed to obtain the second part of the equation [7]. The resulting band structure in the nearest neighbor approximation is plotted in Figure 2.1. It possesses peculiar features, e.g. the saddle point M. Here, the curvature of the dispersion becomes zero and the density of states exhibits a van Hove singularity. While the saddle point is important for the optical properties of graphene [22], the electronic properties are related to remarkable band structure features at the corner points

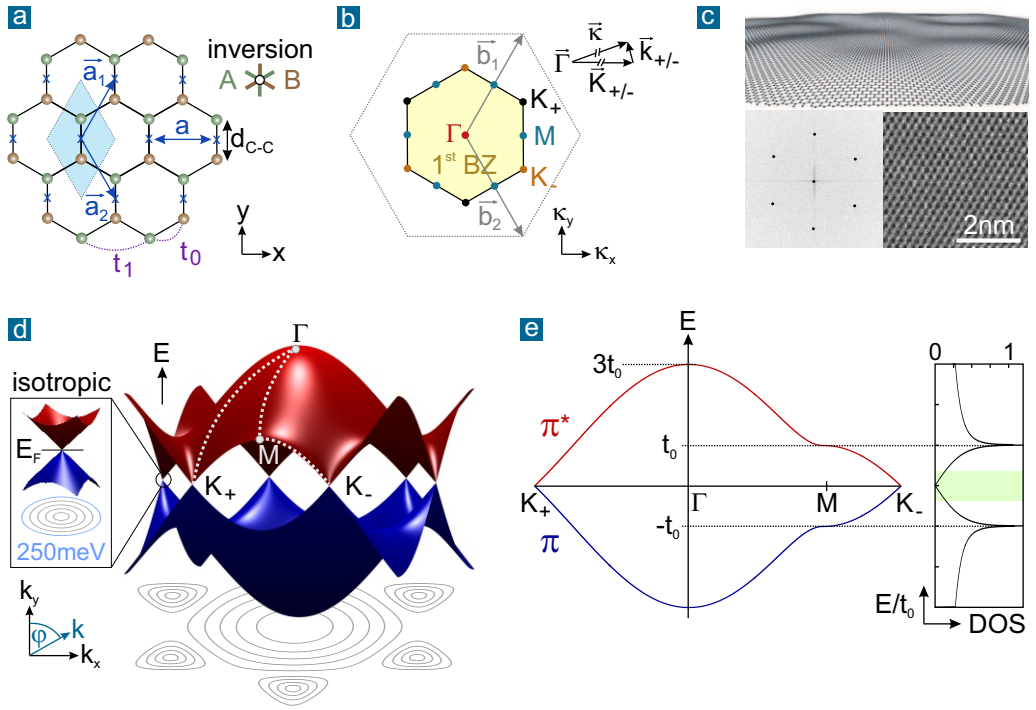

Figure 2.1. Lattice and band structure of graphene. a: The hexagonal lattice in real space with the unit cell containing two carbon atoms which belong to different sublattices A and B that are connected by inversion symmetry. b: Hexagonal first Brillouin zone of graphene with all important symmetry points indicated. Vector diagram for small k vectors around the Dirac points at $K_\pm$. c: Bottom panels show HRTEM and SAED images of one of our CVD grown monolayer graphene samples confirming the hexagonal structure in real and reciprocal space (images taken by Dr. Algara-Siller, University of Ulm). A simulation of corrugated graphene is given in the top panel (taken from Ref. 16). d: Calculated 3D band structure for graphene considering only the nearest neighbor hopping integral t_0. The valence band (blue) and the conduction band (red) touch exactly at the Dirac points for $E_F = 0$meV. A band gap is absent. The inset shows a zoom-in of the band dispersion in k-space in the proximity of the K point. The dispersion is linear and isotropic. Eigenstates are given in cylindrical coordinates by k and φ. e: Band structure along the path drawn as a dashed line in panel d connecting the pivotal points. The calculated density of states diagram diverges at the saddle point M and is linear close to the K points (DOS diagram modified from Ref. 2).

of the Brillouin zone. One can show that the bonding and antibonding π-bands touch at exactly zero energy at the $K_\pm$ points (given by Equation 2.3). In addition, the band dispersion is linear and isotropic for energies that are relevant for electronic properties ($< 0.5\,\mathrm{eV}$) which leads to a cone-like structure. Since each carbon

atom contributes one electron to the π-band, the valence band is completely filled and the conduction band is empty, so that the Fermi energy for ideal graphene (at zero temperature) lies exactly at the $K_\pm$ points ($E_F = 0\,\mathrm{meV}$). Accordingly, graphene can be considered as a zero band gap semiconductor or semimetal. The tight binding band structure and all related features have been verified by ARPES measurements [23, 24].

2.1.4. Dirac-like Hamiltonian for massless chiral fermions

We now focus on the cone-like band structure around the $K_\pm$ points which are also often called valleys following the nomenclature of semiconductors. The center of the reciprocal coordinate system is shifted to one of the $K_\pm$ points by

$$\mathbf{k} = \boldsymbol{\kappa} - \mathbf{K}_\pm \quad \leftrightarrow \quad \boldsymbol{\kappa} = \mathbf{k} + \mathbf{K}_\pm \,. \tag{2.6}$$

The new $\boldsymbol{\kappa}$-vector can be inserted into $f(\boldsymbol{\kappa})$ from which one only keeps the linear terms in $\mathbf{k}$ which is a good approximation close to the $K_\pm$ points. In the nearest neighbor approximation (only considering the transfer integral matrix element H in Equation 2.4), the system is then described by [7]

$$H_\xi = \frac{\sqrt{3}\,a\,t_0}{2\hbar}\,\hbar \begin{pmatrix} 0 & \xi k_x - i k_y \\ \xi k_x + i k_y & 0 \end{pmatrix} = \xi\, v_F\, \boldsymbol{\sigma}^{(*)} \cdot (-i\hbar\boldsymbol{\nabla})\,, \tag{2.7}$$

where $v_F = \sqrt{3}at_0/2\hbar$ is the Fermi velocity and $\boldsymbol{\sigma} = (\sigma_x, \sigma_y)^T$ is the vector of Pauli matrices in two dimensions. Note that for K_-, $\boldsymbol{\sigma}^* = (\sigma_x, -\sigma_y)^T$. This Hamiltonian resembles the two-dimensional Dirac equation for relativistic massless fermions known from particle physics where, however, the speed of light is substituted by the Fermi velocity $v_F \approx 10^6\,\mathrm{m/s}$. It represents the constant group velocity of the charge carriers in graphene. To emphasize this unique feature, the $K_\pm$ points are also referred to as Dirac points and the linear band structure around them Dirac cones. The eigenstates are described by a two-component spinor wave function in analogy to spin-1/2 systems and holes are interpreted as the antiparticles. This is an immediate result of the Dirac equation implying an intrinsic electron-hole symmetry. The two-component spinor $\boldsymbol{\Psi} = (\Psi_A, \Psi_B)^T$ is not related to the physical spin of the charge carrier. It rather has to be interpreted as a pseudospin reflecting the amplitude of the wave function, or in other words the electronic density, located either solely at sublattice A (spin up) or sublattice B (spin down). In graphene, the electronic density is shared equally between the two sublattices and the wave function is given by a linear combination of pseudospin up and down [7]. By introducing polar

coordinates $k = |\mathbf{k}|$ and $\varphi = \arctan k_x/k_y$ the eigenstates can be written as [25, 26]

$$\mathbf{\Psi}_{s,K_\xi}(\mathbf{r}) = \frac{1}{\sqrt{2}} e^{i\mathbf{kr}} \begin{pmatrix} -i\,\xi\, s\, e^{-i\xi\varphi/2} \\ e^{i\xi\varphi/2} \end{pmatrix}, \tag{2.8}$$

where $s = \pm$ corresponds to the conduction (π^*) and valence band (π), respectively. The corresponding eigenenergies are then given by [25, 26]

$$E(\mathbf{k}) = s\,\hbar\, v_F\, |\mathbf{k}|\,. \tag{2.9}$$

This linear equation reflects the massless relativistic character of the charge carriers moving collectively with the same group velocity $v_g = \partial E/\hbar\partial k = v_F$. An unambiguous experimental proof of the linear dispersion is the square root dependence of the cyclotron mass on carrier density (see Figure 2.2c and Refs. 14, 27).

An important parameter characterizing the relative strength of the Coulomb interactions is the so-called r_s-parameter. It relates the typical Coulomb interaction energy, which scales with the inverse of the interparticle distance or $\sqrt{n}$, to the kinetic energy (n is the carrier density). Because to first approximation the dispersion is linear, also the kinetic energy increases as $\sqrt{n}$ (instead of being proportional to n in the parabolic dispersion case), and the r_s-parameter is independent of density. It is only influenced by the dielectric environment (e.g. capping layer κ_1 and substrate κ_2). It can also be interpreted as the effective fine structure constant in graphene α_g and characterizes the interaction between the Dirac fermions (carrier-carrier interactions). It is given by [7, 28–32]

$$r_s(=\alpha_g) = \frac{2}{\kappa_1+\kappa_2}\frac{e^2}{4\pi\epsilon_0\hbar v_F} < 2.2\,(\approx 0.5-0.9 \text{ on hBN or } SiO_2)\,. \tag{2.10}$$

This is in stark contrast to conventional 2DEG with parabolic energy dispersion where $r_s \propto m^* \cdot n^{-0.5}$. In those systems the interactions become dominant at lower densities [7]. The density-independent r_s value in graphene results in an unusual screening behavior of impurities (see Section 2.3.1). Moreover, theoretical studies have extended the analogy to relativistic particles further by predicting a logarithmic renormalization of α_g for $E \to 0$. This effect becomes observable in very clean graphene devices with low extrinsic disorder (see Section 2.1.7).

From the eigenstates it follows that the pseudospin orientation is always coupled to the direction of the $\mathbf{k}$-vector [2, 7]. The expectation value of the pseudospin operator equals

$$\langle\boldsymbol{\sigma}\rangle_{s,\xi} = s\,\xi\,\frac{\mathbf{k}}{|\mathbf{k}|}\,. \tag{2.11}$$

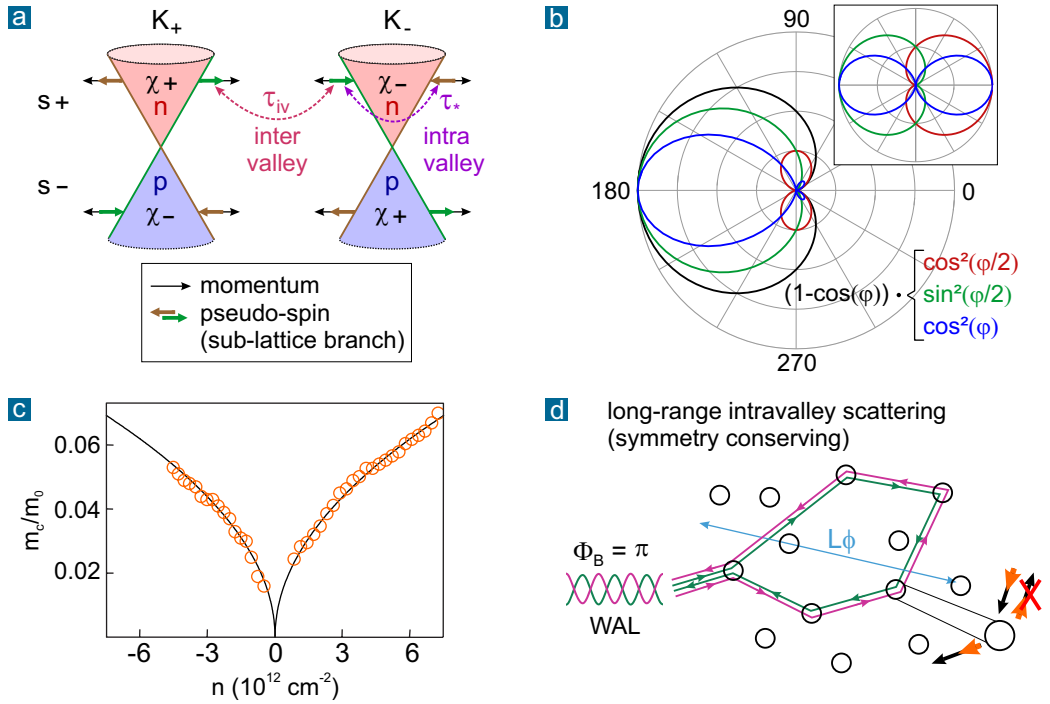

Figure 2.2. Massless chiral carriers in graphene. a: Schematic illustration of the two inequivalent valleys with opposite chirality $\chi = s\xi$. Intravalley direct backscattering is only possible when the disorder can break the A-B symmetry and flip the pseudospin. Intervalley backscattering does not require a pseudospin flip but a large momentum transfer which corresponds to a short range potential in real space. b: The inset shows the angular dependence of the matrix elements $|\langle\Psi(\mathrm{k},\varphi=0)|\Psi(\mathrm{k}',\varphi')\rangle|^2$ for intravalley (red), intervalley or symmetry breaking scattering (green), and for massive chiral carriers in bilayer graphene (blue). The resulting angular dependence of the Boltzmann transport scattering rate according to Equation 2.50 is drawn in the main picture and clearly reveals that for monolayer graphene both small and large angle scattering is suppressed for symmetry conserving long range potentials. c: The square-root density dependence of the cyclotron mass, determined from the temperature dependence of Shubnikov de Haas oscillations, clearly proves the linear dispersion of the charge carriers (taken from Ref. 33). d: When long range intravalley scattering dominates, a single scattering event cannot flip the pseudospin. Then the two counterpropagating closed paths exhibit a net difference of 2π in pseudospin rotation. This corresponds to a Berry phase of π between the two wavefunctions which results in destructive interference of the two paths known as weak anti-localization (WAL). Inelastic scattering with phonons and electrons reduces the dephasing length L_Φ and so the number of possible paths for quantum interference.

As a result, the pseudospin is always parallel or antiparallel to the momentum for one set of valley and band indices, e.g. it is parallel (antiparallel) for the conduction

band of the K_+ (K_-) valley, and vice versa for the valence band (as shown in Figure 2.2a). Hence, electronic states of graphene around the Dirac points have well defined chirality [2]. In summary, the electronic states in graphene host massless *and* chiral quasiparticles. We emphasize that this peculiar characteristic is inherent to the two-dimensional honeycomb lattice structure built from a single atomic species[1].

The pseudospin leads to another property reminiscent of physical spin 1/2. A full rotation of the k-vector and, accordingly, of the pseudospin ($\varphi = 2\pi$) results in a phase change of π in the wave function due to its pseudospin part. This (geometrical) Berry phase $\Phi_B = \pi$ is non-trivial for graphene and has profound implications on its electronic transport [7, 8, 25]. Pseudospin rotation related effects are, among others, the Klein tunneling effect [37], the unusual half-integer quantum Hall effect, the odd shift in Shubnikov de Haas oscillations (see Appendix B.1 and Ref. 8), the suppression of intravalley direct backscattering ($k \to -k$ in a single event) as well as the absence of coherent backscattering (weak antilocalization) if pseudospin conserving long range disorder dominates (as discussed in the following subsection).

2.1.5. Scattering angle restrictions for chiral carriers

The chiral nature of graphene's carriers, reflected by the pseudospin part of its wave function, imposes additional constraints for scattering [7, 8, 25, 26]. Let us first consider the case of direct backscattering in a single scattering process $\mathbf{k}' = -\mathbf{k}$, which in cylindrical coordinates corresponds to $\varphi' = \pi$ assuming in the following without loss of generality $\varphi = 0$. By studying Figure 2.2a it is obvious that (intraband) direct backscattering is only possible for two cases: i) Intravalley direct backscattering requires a potential that can break the A-B sublattice symmetry and thus can flip the pseudospin; ii) Intervalley direct backscattering does not require symmetry breaking but a large momentum in reciprocal space. The latter case (ii) is realized only for short range potentials which act on a range shorter or on the order of the interatomic distance in graphene. The disorder potential must have Fourier components $V(\mathbf{q})$ such that $\mathbf{q} = \mathbf{k} - \mathbf{k}' \approx \mathbf{K}$. Typical examples are vacancies and point defects described by delta potentials (with $V(\mathbf{q}) = \text{const.}$) [25, 26, 32, 38] but also ripples and acoustic phonons in graphene qualify [39]. By definition, this short range disorder can differentiate between the two sublattices and hence leads

1 Bilayer graphene's Hamiltonian describes chiral but massive quasiparticles in two parabolic bands. This results in a trivial Berry phase of $\Phi_B = 2\pi$ and a non-zero backscattering probability $w(\varphi) = \cos^2(\varphi)$ [7, 34, 35]. A single layer of hexagonal boron nitride is two-dimensional and shares the honeycomb structure with graphene. However, it has two chemically inequivalent atoms (B and N) in the unit cell making it a wide band-gap insulator [36].

to scattering in between and within all valleys (and bands), including intra-valley scattering [26]. In contrast, any arbitrary *scalar* long range potential $V(\mathbf{r})$ that acts on a range (much) larger than the interatomic distance can, by definition, not differentiate between the sublattices and thus is unable to influence the pseudospin. Hence, direct backscattering in a single event is completely suppressed: intervalley scattering is not possible because only small momentum $\mathbf{q} \ll \mathbf{K}$ is available for the scattering process and intravalley scattering is not allowed since it cannot alter the pseudospin [25, 26]. We emphasize that this is strictly valid only for *scalar* long range potentials. As we will discuss in section 2.3.2, gauge potentials arising from random strain fluctuations caused by substrate roughness, have long range character. But as a two component vector potential, they can break the sublattice symmetry and flip the pseudospin [18, 40].

The angular dependence of the transition matrix element for a symmetry conserving (intraband $s = s'$) intravalley scattering process from $\mathbf{k}$ to $\mathbf{k}'$ is given by the pseudospin part of the wave function (Equation 2.8)

$$|\left\langle \mathbf{\Psi}_{s,K_+}(k, \varphi = 0)|V|\mathbf{\Psi}_{s,K_+}(k', \varphi')\right\rangle|^2 = |V(k - k')|^2 \cos^2(\varphi'/2)\,. \tag{2.12}$$

This angular dependence $w(\varphi') = \cos^2 \varphi'/2$ is plotted in the inset of Figure 2.2b. Not only is direct backscattering completely suppressed, there is also a general tendency for small angle scattering. As a result, the chiral nature of graphene's carriers allows for very high mobility when the dominating scattering potential is symmetry conserving and long-ranged. Exactly the same argumentation explains why metallic carbon nanotubes (possessing same chiral attributes) show extremely large conductivity [25, 26]. In contrast, the angular dependence of intervalley or symmetry breaking intravalley scattering, e.g. by the gauge potential generated if random strain fluctuations are present [18, 40], is $w(\varphi') = \sin^2 \varphi'/2$ and gives exactly the opposite picture. Direct backscattering is allowed and there is a general tendency towards large angle scattering (inset of Figure 2.2b). To complete the picture, the angular dependence for massive chiral carriers in bilayer graphene with a trivial Berry phase of 2π is added in the graph ($w(\varphi') = \cos^2 \varphi'$) [7, 34, 35].

Another important consequence of the Berry phase π in single layer graphene is the absence of coherent backscattering when pseudospin conserving long range intravalley scattering dominates. The latter implies that the pseudospin cannot be flipped in a single scattering event. Let us consider several of such elastic scattering events that form a closed path (Figure 2.2d). When the dephasing length is long enough, the phases acquired by two carriers traveling in clockwise and counterclockwise direction are well defined. For graphene, the difference in the angle of rotation of the pseudospin for the counterpropagating paths is then 2π [41]. This corresponds to a

phase difference of π in their wave functions (Berry phase) leading to destructive interference. This is reflected by a slight increase of conductivity at lower temperatures when the wavelike nature of carriers becomes relevant. This weak antilocalization (WAL) of the charge carriers is destroyed when a time-reserval symmetry breaking magnetic field is applied. It results in a negative magnetoconductance (positive magnetoresistance) [41]. Note that the WAL effect in graphene is not related to strong spin orbit coupling of the electronic spin itself (as in conventional systems showing WAL) but reflects the chiral nature of its carriers which is ultimately a consequence of its 2D hexagonal lattice. Already bilayer graphene, like other conventional 2DEG without strong SOC, shows coherent backscattering and weak localization resulting in a positive magnetoconductance (negative magnetoresistance). In real devices, quantum interference effects in monolayer graphene are far more complex due to the presence of several elastic scattering mechanisms. Intervalley scattering allows the counterpropagating carriers to occupy different valleys restoring the condition of weak localization [41]. Further, symmetry breaking mechanisms for intravalley scattering can flip the pseudospin and suppress localization effects [18, 42]. There is strong experimental evidence that such a symmetry breaking mechanism prevails at low temperature. The latter is discussed in more detail in Appendix B.1.

2.1.6. Emergence of Moiré superlattices in graphene on hBN

We shortly discuss the effects that can occur when graphene resides on a substrate possessing very similar atomic structure. Here, we focus on hexagonal boron nitride as a substrate[2]. The latter has become the substrate of choice in graphene research allowing the reliable manufacturing of high mobility devices (see Chapter 3). Graphene and this layered insulator share the same planar hexagonal atomic structure. The lattice mismatch is only 1.8% [44]. Under the premise of an atomically clean interface without residues, a Moiré pattern may emerge when graphene is stacked on top of hexagonal boron nitride. The periodicity of this Moiré superlattice depends on the relative rotation angle ϕ between the two crystallographic planes, as shown in STM measurements [44]. The maximum superlattice constant of about 13 nm is reached for $\phi = 0°$ but quickly drops with increasing angle (turning hBN into an insulating substrate causing no changes in the electronic structure of graphene). The weak periodic potential induces spectral reconstructions near the edges of the

2 Moiré patterns also play an important role when two single layers of graphene are rotated slightly with respect to each other. These twisted bilayers are beyond the scope of this work but reveal other interesting band structure reconstructions, e.g. the emergence of low energetic saddle points with van Hove singularities that can be reached by conventional gating [43].

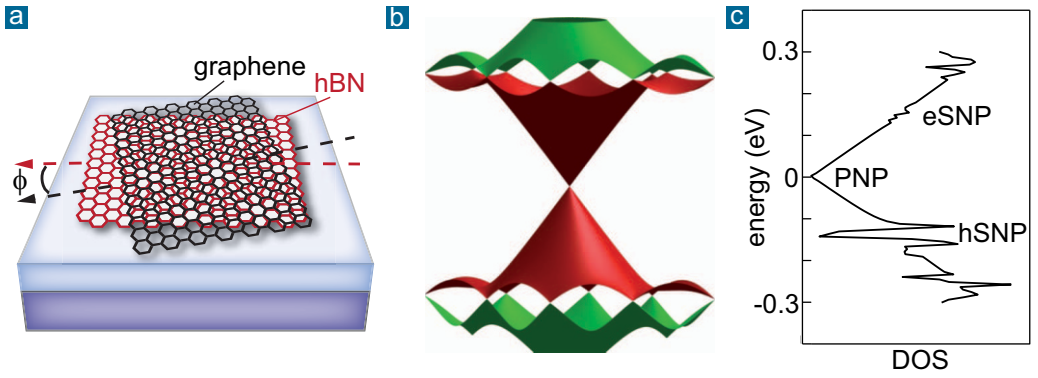

Figure 2.3. Moiré superlattice in graphene on hexagonal boron nitride. a: Moiré pattern emerge when the two crystallographic planes have a small rotation angle ϕ (from Ref. 50). b: 3D band structure around the primary Dirac cone revealing the spectral reconstructions due to the weak periodic potential. Secondary Dirac cones form at the edges of the superlattice Brillouin zone (from Ref. 45). c: Detailed calculation of the density of states reveals asymmetry between the hole-side secondary neutrality point (hSNP) in the valence band and the electron-side secondary neutrality point (eSNP) in the conduction band (from Ref. 46).

superlattice Brillouin zone [45, 46]. Secondary Dirac cones form at around 150-200 meV both in the valence and in the conduction band of graphene (a hole-side secondary neutrality point hSNP, and an electron-side secondary neutrality point eSNP). For the chemical potential to reach these secondary neutrality points, densities of a few $10^{12}\,\mathrm{cm}^{-2}$ are needed. Such densities can be induced by conventional electrostatic gating. The energies of the superlattice Brillouin zone edges, however, increase fast and far beyond electrostatic gating capabilities with decreasing superlattice constant (increasing rotation angle). Hence, practically feasible devices require $\phi < 1°$ [45]. This precision can be achieved either by accurate optical alignment along crystallographic directions (edges of the flakes) or by annealing of arbitrarily stacked but extremely clean hBN-graphene-hBN heterostructures. Without being pinned through trapped contamination and bubbles, the flakes slide on each other on micrometer scale during thermal treatment until graphene self-aligns to one of the hBN surfaces (which corresponds to the local energy minimum) [47, 48]. When graphene is encapsulated between two hBN layers, it is important to ensure that graphene is aligned only to one and misaligned to the other hBN layer such that the second hBN layer does not affect the Moiré superlattice formation [48, 49].

The band dispersion near the secondary neutrality points is linear as for the primary neutrality point (PNP), sometimes also referred to as original neutrality point

(ONP). Accordingly, the characteristic features in the magneto-transport measurements well known from the primary neutrality point are replicated at higher gate voltages (when the chemical potential reaches the SNPs). As shown for our device presented in Figure 4.10 and Section 4.4.2, one observes two satellite resistance peaks which are pronounced at low temperature. Moreover, the sign of the measured Hall coefficients changes at the SNPs. This indicates that electron states are occupied within the valence band of graphene and vice versa which agrees with the theoretical band structure (see Figure 2.3 and Refs. 45, 50). Inherent to the Moiré superlattice is a substantial electron-hole asymmetry between the valence and conduction band. It appears in all electronic properties, including the density of states [46]. Due to this asymmetry, the transport features of the hole-side neutrality point are much more pronounced than for the electron-side (see Figure 4.10). The asymmetry stems from different hopping amplitudes in the Hamiltonian for carbon-above-boron and carbon-above-nitrogen positions, as discussed in Ref. 46. Moreover, the linear dispersion near the secondary Dirac cones is flatter than at the primary Dirac cone. It is characterized by a Fermi velocity v_F^* that is about half of the one at the primary Dirac cone v_F^0 ($v_F^* \approx 0.5 v_F^0$) [36, 44, 45]. Since thermal generation of charge carriers at the neutrality points with linear energy dispersion is proportional to $1/v_F^2$ (see Section 2.2.2), the transport features related to the secondary neutrality points are subject to stronger thermal broadening [45]. As shown in Figure 4.10, the resistance peak at the hole-side is very broad at room temperature (while its equivalent on the electron side cannot be distinguished anymore).

So far, the two rotated lattices were assumed to be rigid and their lattice mismatch to be constant. A more realistic picture is that carbon atoms experience a structural relaxation adapting to the periodicity of the superlattice potential [49, 51]. In other words, the carbon atoms are slightly strained in-plane (with a sign that varies in space) to compensate the lattice mismatch and end up under a certain array of carbon-above-boron and carbon-above-nitrogen positions [46]. Such commensurate relaxation leads to the opening of a band gap at the CNP of several meV which is further enhanced to about 20 meV by electron-electron interactions [51]. The band gap is reflected in insulating, thermally activated transport behavior at the CNP [47, 49, 52, 53]. The emergence of the SNPs is a more robust effect since it is linked to the formation of a Moiré pattern alone (with or without lattice mismatch) [36]. The commensurate relaxation, however, requires an extremely clean interface and rotational alignment precision. Incommensurability is much more likely achieved, e.g. due to residual impurities that may act as pinning sites, to an increased interface roughness causing additional in-plane strain, or to another aligned hBN layer distorting the superlattice [36, 49, 51]. All in all, despite the substantial changes in the electronic structure, graphene's carrier mobility is not degraded by the Moiré pattern or the commensurate relaxation [46, 51].

For the majority of our devices presented in this work, graphene was not purposely aligned to the encapsulating hBN flakes. Instead, hexagonal boron nitride simply serves us as ideal insulating substrate which improves substantially the electrical properties of graphene but does not alter its electronic structure. Accordingly, Moiré pattern effects and commensurate relaxation related gap opening can be neglected in our further discussion besides for the intentionally aligned heterostructure we present separately in Section 4.4.2.

2.1.7. Many body interactions: velocity renormalization and scattering

In this section we give a brief overview on the effects caused by many body interactions in graphene. At low energies, long-range Coulomb interactions between the charge carriers induce considerable modifications of the linear (single-particle) Dirac spectrum. As discussed below in the first subsection, this is a direct consequence of the logarithmic renormalization of the Coulomb interaction constant in graphene α_g (also referred to as graphene's effective fine structure constant, in analogy to the renormalization of the fine structure constant in the quantum electrodynamic theory) [4, 54]. Moreover, carrier-carrier interactions not only renormalize the quasiparticle properties, such as the quasiparticle velocity (equivalent to the modified band dispersion). They also play a dominant role in carrier transport when their collision rate becomes the largest in the system (which can be realized in recent graphene devices with extremely low extrinsic disorder). This gives rise to a third regime of carrier transport, aside from the common diffusive and ballistic transport regimes which are well understood in a single-particle picture. Instead of behaving like non-interacting quasiparticles, collective phenomena and fluid-like behavior of the carriers characterize this third regime. Therefore, hydrodynamic quantities are introduced to describe the viscous transport. The second subsection of this section discusses briefly the two different regimes of carrier-carrier scattering and its temperature and density dependence. With this background, it can be estimated under which conditions many body effects can play a role in transport measurements in suitable devices. The hydrodynamic transport regime and the key experiments revealing viscous transport features are then reviewed in Section 2.4.

Spectrum and Fermi velocity renormalization

We have shown in Sections 2.1.3 and 2.1.4 that, in nearest neighbor approximation, the tight binding model predicts a linear energy dispersion at the vicinity of the

Dirac points. As a consequence, the Fermi velocity is a constant and determined solely by the lattice constant and the nearest neighbor hopping energy (see Equation 2.7). Its value[3] is then $v_F^0 \approx 1 \cdot 10^6\,\mathrm{m/s}$ which was confirmed by early transport measurements based on the analysis of cyclotron resonances and Shubnikov-de-Haas oscillations [21, 33]. These experiments were conducted at carrier densities beyond $10^{12}\,\mathrm{cm}^{-2}$ and the density was tuned only over a small interval (not over several orders of magnitude). At such high densities, however, Coulomb interactions are efficiently self-screened by the charge carriers in graphene. In addition, renormalization effects that scale logarithmically are hard to detect when the density (the energy) is not varied over a very wide range. As we will elaborate extensively in Section 2.2 and Chapter 3, extrinsic disorder in real devices determines the lower energy limit to which the Dirac cone physics can be accessed experimentally. When the average density is tuned to zero, the system breaks up into a spatially inhomogeneous landscape of electron and hole puddles. This means that, locally, the system is still doped. Since the amplitude of these puddles is directly linked to the strength of the extrinsic disorder, the latter has to be reduced substantially in order to effectively approach the Dirac point (the low-energy regime, $E \to 0$, in the band structure). Then, the spectral changes caused by many body interactions become important and observable.

Let us now assume that extrinsic disorder is completely absent. When the Dirac point ($E \to 0$) is approached, long-range Coulomb interactions cannot be screened efficiently any longer since the number of charge carriers decreases. Taking into account self-screening effects the screening behavior becomes highly complex and, ultimately, the screening length diverges exactly at the Dirac point (also referred to as charge neutrality point) [56]. In case of such theoretical divergences, the Landau theory fails and carrier-carrier interactions have to be taken into account by using the renormalization group theory which solves the divergence problem by renormalization of the coupling constant, in full analogy to quantum electrodynamics (QED) [54, 56]. For small coupling (interaction strength), the self-energy corrections are calculated by standard (first order) perturbation theory [4]. However, the Coulomb interaction constant in graphene α_g^0 is not small (within the meaning of $\alpha_g^0 \ll 1$). In Section 2.1.4 we have derived α_g^0 in case of linear energy dispersion

3 We use the superscript 0 in the notations of v_F^0 and α_g^0 to indicate that these values are related to the linear energy dispersion (in the tight binding model). Note, at the ultraviolet cutoff, the corresponding quantities in the renormalization group theory converge to these values. The cutoff energy defines the energy up to which the spectrum is Dirac-like (linear). It is typically assumed to be $\approx 3\,\mathrm{eV}$ in the calculations [4, 55, 56].

resulting from the (single-particle) tight binding model (see Equation 2.10). Introducing the background dielectric constant $\kappa = (\kappa_1 + \kappa_2)/2$ and the speed of light c, one can rewrite the expression in the following manner [12]

$$\alpha_g^0 = \alpha_{QED} \frac{c}{v_F^0 \kappa} = \frac{1}{137} \cdot \frac{300}{\kappa} \approx 2.2/\kappa \,. \tag{2.13}$$

Accordingly, α_g^0 is on the order of one and the employment of first order perturbation theory is conceptually critical, even though first calculations reproduced well the experimentally observed logarithmic enhancement of the Fermi velocity for very low carrier densities (which is discussed in detail below) [4, 54, 56]. There is an ongoing discussion[4] among theorists about the right theoretical approach to resolve this issue since it was also shown that taking into account the second order in the perturbation theory already produces wrong results (disagreeing with the experiments) [59]. We restrict us here to the random phase approximation theory as described in Ref.58.

Under real conditions, graphene samples can be tuned to very low energies (densities) but, owing to extrinsic disorder, will never reach exactly the critical Dirac point. Hence, graphene remains always doped (extrinsic, in the language of theorists). Among others, the group of Das Sarma has shown in a series of publications [57–59] that the dynamical and static screening of electron-electron, electron-lattice, and electron-impurity interaction by the charge carriers themselves is described well within a self-consistent random-phase approximation (RPA) theory (see Section 2.2.3). Moreover, the application of RPA in perturbation theory results in rapidly converging next-to-leading order corrections [59]. For excitations at the Fermi level (for $k = k_F$), the leading-order RPA result for the renormalization of the Fermi velocity can be given in an analytical form [55, 57, 58]. It reads

$$\frac{v_F^*(n)}{v_F^0} = 1 - \frac{\alpha_g^0}{\pi}\left[\frac{5}{3} + \ln(\alpha_g^0)\right] + \frac{\alpha_g^0}{8}\ln\left(\frac{n_c}{n}\right) , \tag{2.14}$$

where n_c is the charge carrier density corresponding to the ultraviolet cutoff energy ($\approx 3\,eV$) and n is the charge carrier density which is typically controlled via a back gate in experiment. The RPA theory predicts a logarithmic enhancement of the Fermi velocity at small carrier densities (like the bare perturbation theory). Moreover, the renormalization effect is expected to be stronger for smaller dielectric background constants κ (larger α_g^0 values). Hence, the strongest effects should be observable for suspended graphene in vacuum when the dielectric screening of

4 We refer to References [57–62] for the discussion about alternative theories and improvements, such as a nonperturbative approach [60, 61], the combination of random phase approximation theory and perturbation theory [57–59], or even detailed Hartree-Fock calculations on the honeycomb lattice [62].

Coulomb interactions is minimal. We emphasize that Equation 2.14 describes the renormalized electronic spectrum of graphene at a certain gate controlled density n for excitations at the Fermi level (at $k = k_F \propto \sqrt{n}$). This corresponds to electrical transport measurements where only carriers near $k = k_F$ contribute. Therefore, transport measurements can reveal only the density dependence of the renormalized spectrum since the probed momentum (or energy) range is set by the density (and cannot be varied independently).
However, in the general case, the spectral renormalization caused by many body interactions is a function of both density and momentum (independently). In addition, since α_g^0 in graphene is independent of the carrier density, the dispersion renormalization also depends additionally on the dielectric environment κ. All in all, the general renormalized spectrum is then a function of (n,k,κ) which cannot be expressed anymore in an analytical form [55, 58]. Instead, a numerical approach is proposed and presented in Ref. 58.

The theoretical predictions have been verified by a multitude of experiments employing a variety of techniques, such as infrared spectroscopy [65], scanning tunneling Landau level spectroscopy [55, 66, 67], (magneto-)transport [56, 68], APRES [63, 69, 70], or quantum capacitance measurements [71]. In particular, the results could not be explained quantitatively without taking into account the enhancement of the Fermi velocity at low densities (that means carrier-carrier interactions). In the following, we want to discuss three important experiments whose main findings are replotted in Figure 2.4.

Elias and coworkers (Ref. 56) repeated the analysis of the decay rate of quantum oscillations which had already been conducted on low quality devices (as mentioned above). In contrast, their measurements were performed on suspended graphene with extremely low disorder. The Dirac point could be approached within a few millielectronvolts and the density could be tuned over a wide range and to very low values (without formation of electron-hole puddles). Moreover, since dielectric screening is absent for suspended devices, the renormalization effects are particularly strong. As shown in panel a of Figure 2.4, their measurements provide clear evidence for the logarithmic divergence in the Fermi velocity as the density is tuned to very low values. This is in good quantitative agreement with the theoretical prediction given in Equation 2.14 (for excitations at $k = k_F$). The electronic spectrum changes considerably compared to the linear dispersion (having an hourglass shape as sketched in the inset of panel a). In particular, carrier-carrier interactions reduce the density of states at low energy. An important question, to our knowledge not considered neither in theory nor specifically in experiments, is how the thermal activation of charge carriers near the charge neutrality point (Dirac point) is quantitatively changed in case of the renormalized spectrum (see also Section 2.2.2 on

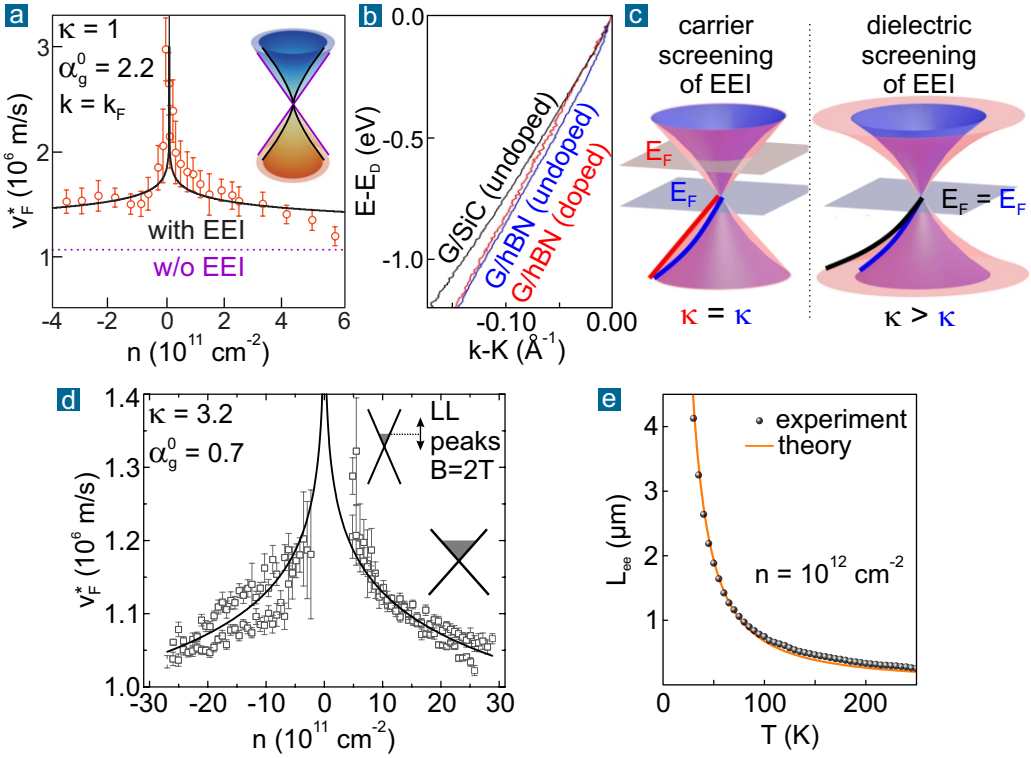

Figure 2.4. Fermi velocity renormalization by electron electron interactions (EEI). a: Logarithmic divergence in the Fermi velocity at low carrier densities obtained by the analysis of the decay rate of SdHO (from Ref. 56). Note, transport measurements only probe the renormalized spectral function at $k = k_F$. The inset shows a sketch of the modified electronic spectrum near the Dirac point when EEI are considered (modified from Ref. 56). b: ARPES measurements probing the energy-dependence of spectral changes for almost undoped graphene on SiC (larger κ), undoped graphene on hBN (smaller κ), and doped graphene on hBN. c: Interpretation of the data plotted in panel b (same line colors). The electronic spectra for different screening types and doping levels are given (see main text). Panels b,c were modified from Ref. 63. d: Renormalized Fermi velocity for graphene on hBN, derived from Landau level tunneling spectroscopy for higher densities than in panel a (modified from Ref. 55). For each density, tunneling spectroscopy of the Landau levels allows probing the k-space away from k_F. The study finds the dispersion to be always linear, however with increasing slope (Fermi velocity) as the density is lowered. e: Temperature dependence of the characteristic length scale of electron electron interactions $L_{e-e} = v_F \tau_{e-e}$ for doped graphene in the Fermi liquid regime, as discussed in the second subsection (from Ref. 64).

thermal generation of charge carriers). We will come back to this problem in Section 4.3 when analyzing our Hall measurements.

As aforementioned, transport measurements lack the possibility to tune the momentum (or energy) independently from the density. Hence, they cannot probe the full parameter space of the renormalized spectral function. This issue can be addressed by angle-resolved photoemission spectroscopy (ARPES) measurements as conducted by Siegel et al. in Ref. 63. Panel b of Figure 2.4 presents APRES traces for three different devices. The effect of carrier screening on the carrier-carrier interactions becomes obvious when undoped and doped graphene on hBN devices are compared (blue and red lines, respectively). The dispersion in case of doped graphene is linear (here, density n is fixed but momentum k is varied). In contrast, the electronic spectrum is clearly renormalized in case of undoped graphene near the Dirac point (proving the sketched spectrum in the inset of panel a). Moreover, undoped samples with different dielectric environment are compared (blue and black lines, respectively). Both traces are non-linear but differ at high momentum values. This reveals the effect of the dielectric screening on carrier-carrier interactions. The modifications of the Dirac cone for different screening types is summarized in panel c.

The devices used for the ARPES study were not equipped with a back gate. The latter allows an arbitrary and wide tuning of the density. In their experiments using scanning tunneling spectroscopy of Landau levels, Chae and coworkers (Ref. 55) combine the handiness of a back gate (to tune the density like Elias et al.) with the benefit of their technique that also energies away from the Fermi energy can be examined (instead of right at the latter as in conventional transport measurements). Their main result is reproduced in panel d of Figure 2.4. For each density they acquire the Landau level spectrum from which one can extract the Fermi velocity and gain information on the shape of the dispersion for an energy range of $\pm 200\,\mathrm{meV}$. Within the examined density range, their data clearly shows that the Fermi velocity is enhanced for lower densities (in agreement with the experiment of Elias et al.). Moreover, their analysis suggests that the renormalized dispersion remains linear at all tested densities but the slope (the Fermi velocity) increases. They call this "squeezing of the Dirac cone". This conforms with the theoretical predictions discussed earlier, since their experiment is conducted at comparatively high densities (larger than the range examined by Elias et al.). At such high densities, carrier-carrier interactions are efficiently screened by the high number of charge carriers and the spectrum remains linear with a slightly modified slope (compare also to panel c of Figure 2.4).

Carrier-carrier scattering regimes and their time scale

The carrier-carrier interactions not only modify the band dispersion but also manifest in characteristic hydrodynamic transport phenomena[5]. As discussed in detail in Section 2.4, hydrodynamic effects dominate when the carrier-carrier scattering time is the smallest in the system (or, equivalently, when the mean free path for carrier-carrier collisions is the shortest length scale). Graphene offers important prerequisites in order to meet this condition. The electron-phonon coupling is exceptionally weak and intrinsic disorder (such as lattice defects) is very low. Momentum-relaxing scattering is typically dominated by extrinsic disorder over a wide parameter space (regarding density and temperature). The latter can then be minimized by proper device processing (see Chapter 3).
In order to estimate whether the carrier-carrier collision length scale can ever[6] become the smallest in a given graphene device, knowledge of the temperature and density (or chemical potential μ) dependence of the carrier-carrier collisions is crucial.

One has to distinguish between two different regimes of carrier-carrier scattering in graphene. At high density, when the chemical potential is much larger than the thermal energy ($\mu \gg k_B T$), the possible scattering pathways are restricted to states close to the Fermi surface [12]. This regime is referred to as Fermi liquid (FL). Only electrons (or holes) are present and, accordingly, only scattering between electrons (or holes) has to be considered. However, at low densities, when the chemical potential approaches the Dirac point becoming much smaller than the thermal energy ($\mu \ll k_B T$), the Fermi surface vanishes and the former argumentation no longer applies. Moreover, electrons and holes coexist due to thermal activation at finite temperature and also scattering between electrons and holes is possible. Since the screening of Coulomb interactions becomes ineffective, a strongly interacting

5 Here, we focus on viscous effects seen in (magneto-)transport (in nonquantizing magnetic fields). Carrier-carrier interactions are also important to understand unconventional (fractional) quantum Hall states [72–74], graphene plasmonics [75], and the emergence of parabolic magnetoresistance at high densities [76]. They are also claimed to be the dominant contribution to the dephasing time in quantum interference experiments (see Ref. 77 and Appendix B.1). Moreover, the strength of electron-phonon coupling is also renormalized, in particular for optical phonons near the charge neutrality point (see Ref. 78 and Section 2.3.2).

6 Each device is characterized by its device dimensions (e.g. the flake size in the order of few micrometer) and by a certain level of extrinsic disorder. As a consequence, carrier-carrier collisions will compete with three different types of momentum relaxation scattering: scattering at extrinsic disorder, boundary scattering (in the ballistic case at low temperature), and phonon scattering (at elevated temperatures). Hence, the so-called hydrodynamic window in the parameter space by density and temperature is not universal and only realized in devices of extremely high quality (see Section 2.4).

electron-hole "plasma", also referred to as Dirac fluid regime (DF), forms [12, 79]. The carrier-carrier scattering times are then given by [12]

$$\tau_{c-c}^{FL} \propto \frac{1}{\alpha_g^2}\frac{\hbar\mu}{k_B^2 T^2} \qquad \tau_{c-c}^{DF} \propto \frac{1}{\alpha_g^2}\frac{\hbar}{k_B T}\,, \tag{2.15}$$

from which the characteristic length scales can be calculated ($L_{c-c} = v_F\,\tau_{c-c}$). The Fermi liquid regime can be accessed unambiguously in experiments and the above given scattering time dependence on temperature (and density, not shown) has been verified (as shown in panel e of Figure 2.4, obtained from superballistic transport measurements). The length scale decreases with temperature but increases with density (due to the stronger screening). Hydrodynamic effects in the Fermi liquid regime are thus expected to be observable at intermediate temperatures and densities in sufficiently clean samples with low disorder.

On the contrary, the experimental realization of a pristine Dirac fluid regime is yet under debate, at least for monolayer graphene (see Refs. 79–81 and Section 2.4). As aforementioned, at the charge neutrality point, extrinsic disorder induces electron hole puddles in real devices. The puddles then constitute small patches of Fermi liquid masking the Dirac fluid regime [12, 57]. Only at higher temperatures, when thermally activated carriers smooth out the spatial inhomogeneity, the Dirac fluid regime might recover (if phonon scattering does not yet prevail).

Finally, we refer to Section 2.4 for further treatment of the hydrodynamic regime and the key experiments.

2.2. Charge carriers in graphene

As shown in the previous section, the valence and conduction bands touch at the Dirac point ($E_F = 0\,\text{meV}$). In pristine graphene at zero temperature, the valence band is completely filled and the conduction band is completely empty. No free charge carriers, neither electrons nor holes, are present. Due to the absence of a band gap, this ideal case of intrinsic graphene cannot be achieved in real devices where both finite temperature and external disorder can easily induce charge carriers. The Dirac-like Hamiltonian describing the linear Dirac cones, however, ensures electron-hole symmetry. A finite temperature will activate equal densities of electrons (n_e) and holes (n_h) at the Dirac point, resulting in a finite *total* charge carrier density $n_t = n_e + n_h$ but zero *average* charge carrier density $n = n_e - n_h$. Therefore, this point is also called charge neutrality point (CNP). The Fermi level or, for finite

temperature, the chemical potential μ (with $\mu(T \to 0) = E_F$) can be tuned across the CNP over a wide window by electrostatic gating [1]. This corresponds to an injection or depletion of electrons and thus the average charge carrier density increases or decreases, respectively[7]. Gating, finite temperature, and disorder collude in real devices as we will show in this section.

2.2.1. Ambipolar tuning of the carrier density by field effect gating

For conventional 2D electron systems with a parabolic dispersion the density of states is constant as a function of energy. In stark contrast, the linear dispersion of graphene causes the density of states to depend linearly on energy. Accounting for spin and valley degeneracies in graphene, the DOS is given by [2, 3, 82]

$$\mathrm{DOS}(E) = \frac{g_s g_v}{2\pi(\hbar v_F)^2}\,|E| = \frac{2}{\pi(\hbar v_F)^2}\,|E| = D_0\,|E|\,, \tag{2.16}$$

with the prefactor $D_0 = 1.47 \cdot 10^{14}\mathrm{cm}^{-2}\mathrm{eV}^{-2}$ for $v_F = 1 \cdot 10^6\,\mathrm{m/s}$. We first consider the simple case of ideal (disorder-free) graphene at zero temperature, then the charge carrier density is connected to the Fermi energy E_F according to

$$n = \int_0^{E_F} \mathrm{DOS}(E)\,dE = \frac{E_F^2}{\pi\hbar^2 v_F^2}\,, \tag{2.17}$$

with $E_F > 0$ for electrons and $E_F < 0$ for holes. In experiment, these densities can be tuned at the turn of a voltage knob. Graphene is typically placed on an insulating substrate, e.g. thermally grown SiO_2 on highly doped Si, optionally with an additional thin flake of hexagonal boron nitride between graphene and SiO_2. The oxide thickness is chosen to be about $d \approx 290\,\mathrm{nm}$ in order to enhance the optical contrast of graphene [83]. The system then constitutes a parallel plate capacitor and the gate induced (average) charge carrier density n_g can be calculated for an

7 In this thesis, the notation n_e (or n_h) always refers to the spatially independent or average total electron (hole) density as if the system was spatially homogeneous. In case of disorder-free graphene, the system is obviously homogeneous. As we will discuss in Section 2.2.3, extrinsic disorder can dominate near the CNP and impose a spatially inhomogeneous landscape of locally varying electron and hole densities (which then exhibit a lateral, e.g. (x,y), dependence). However, also in the latter case, the measured quantities in electrical transport, such as the resistivity or the Hall coefficient, can be described adequately by an assumed spatially homogeneous model system. Then, using the two carrier type Drude model, the spatially independent or average electron (n_e) and hole (n_h) densities are determined by fits to the experimental results (see Sections 2.2.4 and 2.3.3).

applied gate voltage V_g using the known relative permittivity for SiO_2 ($\epsilon_r = 3.9$) as follows

$$n_g = \frac{\epsilon_0 \epsilon_r}{e\,d} V_g \equiv \beta V_g \;\; [= \beta(V_g - V_{CNP})]\,. \tag{2.18}$$

Under the reasonable assumption that quantum capacitance effects can be neglected (see Section 2.2.6), this gate induced carrier density is equal to the average carrier density in graphene

$$n_g = n = n_e - n_h\,. \tag{2.19}$$

For disorder-free graphene at zero temperature, considered here, n_g is just either the electron n_e or hole density n_h. A positive gate voltage corresponds to inducing electrons and a negative gate voltage to holes. The gate factor β for the aforementioned parameters is $7.4 \cdot 10^{10}\,cm^{-2}/V$. When a hBN flake is employed as additional substrate, the effective gate factor can be derived from Hall measurements at higher densities and gives the equivalent oxide thickness (EOT). Equation 2.17 can be rewritten to yield the Fermi energy for a given electron or hole density

$$E_F = \mathrm{sign}(n)\,\hbar v_F \sqrt{\pi\,|n|} = \hbar v_F k_F \quad \leftrightarrow \quad k_F = \mathrm{sign}(n)\,\sqrt{\pi\,|n|}\,. \tag{2.20}$$

The above formulas remain a good approximation at finite temperature as long as $k_B T \ll E_F$ holds (e.g. at low temperature and/or high charge carrier density).

The dielectric breakdown field limits the range how far the carrier density in graphene can be tuned by field effect gating. For SiO_2, 10 MV/cm is often stated in the literature. However, this value is based on well engineered gate designs and clean CMOS processing steps. According to our experience, this value needs to be reduced by at least a factor of five, restricting conventional electrostatic gating to a carrier density slightly below $10^{13}\,cm^{-2}$ ($E_F = 370\,meV$). Moreover, to minimize hysteric effects (which will be discussed in Chapter 3), the rule of thumb is to stay below 30 V for 300 nm EOT (1 MV/cm). Hence, the density tunability is limited to $\approx \pm 10^{12}\,cm^{-2}$ ($E_F \approx \pm 100\,meV$). Much higher doping levels can be reached with electrolyte gating [84–87]. Lithium-based liquid or solid electrolytes are typically employed where the lithium ions form an atomically thin electric double layer on the surface of graphene, thereby transferring a large amount of electrons to graphene. For several layers thick 2D materials, Li ions can even intercalate between the layers, causing extremely high doping levels beyond $10^{14}\,cm^{-2}$ ($E_F > 1\,eV$) as shown in bilayer [88] and in hBN-encapsulated monolayer graphene [89].

Aside from such intentional doping, charge transfer between graphene and physical adsorbates or trapped charges in the dielectric occurs also unintentionally. Being

an all-surface material, an initial (average) doping level Δn_i is observed in any non-encapsulated device. As a result, a non-zero gate voltage value V_{CNP} needs to be applied in order to reach charge neutrality and compensate this initial doping (included in the last part of Equation 2.18). Depending on the exact processing steps and the environment, both initial hole doping ($V_{CNP} > 0$) and initial electron doping ($V_{CNP} < 0$) have been observed [90–95]. Graphene devices on SiO_2, however, are typically found to be hole-doped on the order of some $10^{11}\,cm^{-2}$ (for common processing steps), most likely due to moisture-assisted physical adsorption of oxygen molecules and/or silanol groups terminating the SiO_2 surface (see Chapter 3 and Ref. 94, 95).

2.2.2. Thermal generation of charge carriers

We now address the thermal generation of charge carriers in disorder-free graphene by introducing the Fermi Dirac distribution. For a given chemical potential μ and temperature T, the distribution reads

$$f(E) = \frac{1}{e^{(E-\mu)/(k_B T)} + 1} \quad \text{with} \quad \mu(T \to 0) = E_F \,. \tag{2.21}$$

Then the electron density is given by

$$n_e = \int_0^\infty DOS(E)\, f(E)\, dE \,. \tag{2.22}$$

The hole density n_h is calculated accordingly by substituting $1 - f(E)$ and adapting the integration range to negative energies. The chemical potential is determined by considering charge conservation. It can be derived numerically so that the average carrier density is equal to the gate voltage induced density

$$n_e(\mu, T) - n_h(\mu, T) \stackrel{!}{=} n_e(E_F, 0) - n_h(E_F, 0) = n_g = \beta(V_g - V_{CNP}) \,. \tag{2.23}$$

Figure 2.5d shows the dependence of the chemical potential at finite temperature on the gate voltage controlled average carrier density. Of special interest is the charge neutrality point $\mu = 0$ where $n = n_e - n_h = n_g = 0$ and $n_t = n_e + n_h = 2n_e$ has a finite minimum (at finite temperature). In the following we use the notation $n_{t,0}$ for the (residual) total carrier density at the CNP[8]. It can be calculated analytically for the linear dispersion and reads

$$n_{t,0} \equiv n_t(\mu = 0, T) = 2\, n_e(\mu = 0, T) = \frac{\pi k_B^2}{3\hbar^2 v_F^2} T^2 \,. \tag{2.24}$$

8 Here, we consider only thermally activated carriers that contribute to $n_{t,0}$. In the presence of disorder, $n_{t,0}$ is the residual carrier density at the CNP as if the system was homogeneous and both disorder induced carriers and thermally activated carriers contribute.

Assuming $v_F = 1 \cdot 10^6\,m/s$, the total carrier density at the CNP through thermal activation is $2.9 \cdot 10^7\,cm^{-2}$ for T = 4K and $1.6 \cdot 10^{11}\,cm^{-2}$ for T = 300K (in the case of disorder-free graphene).
The situation becomes much more difficult when one includes carrier-carrier interactions (connecting to Section 2.1.7). For example, assuming the gate defines the average density to be $10^{11}\,cm^{-2}$, then Coulomb interactions between the carriers are partially screened. Several experiments suggest the band dispersion to be still linear but its slope, the Fermi velocity, is increased compared to the bare value $v_F^0 \approx 1 \cdot 10^6\,m/s$ obtained from the tight binding model (see Section 2.1.4). In a simplified picture, this suppresses the thermal carrier generation by a factor $(v_F^0/v_F^{norm})^2$ (see Figure 2.5e and Chapter 4 for further discussion).

2.2.3. Disorder induced spatial density inhomogeneities

In real devices, extrinsic disorder, such as charged impurities or random strain fluctuations, cannot be avoided entirely (see Section 2.2.5 and Chapter 3 for details). The extrinsic disorder is randomly distributed and of long range character[9]. Due to the absence of a band gap, such disorder potentials can easily induce charge carriers in graphene, as described in the following.

The long range character implies that the effect of the disorder potential prevails at low carrier densities (near the CNP), because then the screening of Coulomb interactions is weak (as already discussed in the case of carrier-carrier interactions in Section 2.1.7). At higher carrier densities, the screening is efficient and the strength of the disorder potential is reduced. Moreover, the random microscopic configuration of the disorder leads to a *spatially* fluctuating electrostatic (disorder) potential V(x, y) (assuming the graphene layer to be in the (x,y)-plane). Since the electrochemical potential μ_{el} (the global Fermi level) of the system is a constant independent of space (in equilibrium), the chemical potential μ varies equally in magnitude but opposite in sign with respect to the electrostatic (disorder) potential, such that $\mu_{el} = \mu(x, y) + V(x, y) =$ constant [96, 97]. A schematic representation of this correlation is given in panel b of Figure 2.5 assuming the system to be at the CNP

9 Weak short range disorder that can be treated in the Born approximation cannot explain experimental findings such as the linear dependence of the conductivity on the density near the CNP and the observed characteristics of quantum interference effects. Resonant (short range) scatterers are strong enough to add extra structure in the DOS and can explain the linear dependence. However, there are many good reasons (listed in Ref. 7) why such midgap states are not a proper model for graphene devices on conventional substrates. Only for devices with a large amount of artificially generated lattice vacancies (causing a strong modification in graphene's fundamental chemical structure), the picture of resonant scatterers is applicable [3].

($\mu_{el} = 0$). Then, the disorder potential fluctuations are most pronounced both in absolute magnitude (weak screening) and in terms of their relative impact (the disorder induced density fluctuations exceed the average density). The spatial fluctuations in the chemical potential are equivalent to local shifts of the Dirac point energy and to the formation of electron-rich and hole-rich regions, also referred to as electron and hole puddles [96].

In summary, under the presence of extrinsic disorder, the system breaks up into a *spatially* inhomogeneous landscape with fluctuating carrier density when the charge neutrality point is approached (by applying an external gate voltage $V_g = V_{CNP}$ such that $\mu_{el} = 0$ and the gate controlled *average* density n_g is zero). Charge neutrality then corresponds to the system containing an equal total number of electrons and holes by integrating the density over the entire spatial density profile. We emphasize that the spatial inhomogeneity is the main difference to the coexistence of electrons and holes by thermal generation at the CNP in disorder-free graphene, where at any given spatial position both electrons *and* holes are present forming a homogeneous system.

The spatial inhomogeneity is most pronounced at low temperature ($T \to 0$) and near the CNP. Increasing temperature will lead to thermal generation[10] of charge carriers smoothing out the disorder induced profile (as indicated by the Fermi-Dirac distribution in Figure 2.5b). The system can be considered homogeneous at the CNP when the temperature exceeds the disorder equivalent temperature scale (see Table 2.1). The second option to make the system homogeneous is to tune the average density away from the CNP. This corresponds to raising or lowering the electrochemical potential (the global Fermi level). Then, the density fluctuations become negligible compared to the average density and the system is homogeneously electron *or* hole doped[11]. Moreover, an increase in the average carrier density will

10 The thermal generation of charge carriers in a system with spatially fluctuating chemical potential is a highly complex task for theory (which has not been addressed so far, to our knowledge). For instance, there are areas with large (local) μ, where the application of the Fermi Dirac distribution does not lead to a substantial change in the carrier density. On the contrary, there are also areas where μ is (close to) zero and both electrons and holes are thermally generated. Moreover, we know from Section 2.1.7 that the band dispersion (the Fermi velocity) is renormalized when $\mu \to 0$. Hence, there are areas (when considering each puddle for itself) with varying Fermi velocity. Of course, a proper theory has to consider the inhomogeneous system as a whole, e.g. thermally activated hole states can be occupied by nearby electrons, but the above picture focusing on each puddle illustrates the complexity of the problem. In the fourth chapter (Section 4.3), we will discuss this problem in more detail.

11 In other words, the fluctuations of the chemical potential are small compared to the global offset of the chemical potential from zero (the energy fluctuations are small compared to the global energy). Then, for instance, the system is electron doped with areas of slightly less or slightly more electrons. Ultimately, these variations become negligible.

further lead to an enhanced screening of the disorder potential and the size and the amplitude of the puddles will reduce [31, 98].

In order to make quantitative theoretical predictions on the puddle formation, a good understanding of the dominant extrinsic disorder mechanism is needed (as discussed in detail in Section 2.2.5). In most theoretical studies, charged impurities are considered as dominant mechanism. Moreover, the screening of Coulomb interactions by the surrounding charge carriers in graphene has to be modeled properly. In case of electron hole puddles, a self-consistent theory is required: A strong disorder potential will induce puddles with larger amplitude which in turn will more effectively screen this disorder potential. In the following two paragraphs, a general introduction to the random phase approximation (RPA) theory describing the screening of Coulomb interactions and to a self-consistent approximation (SCA) ansatz, which attempts to describe the inhomogeneous system by a homogeneous model system, is given. The amplitude and length scale of the electron and hole puddles are then determined by the strength of the disorder potential and the density-dependent dielectric RPA-screening function (which both vary spatially along with the density fluctuations) [31].

RPA screening theory and puddle formation in a self-consistent picture

The screening of Coulomb interactions in graphene is crucial to understand carrier-impurity, carrier-carrier, and carrier-phonon interactions. It is described in good quantitative agreement to various experiments by the random phase approximation (RPA) theory [3, 31, 57, 59, 99, 100]. The implementation of the correct screening model is therefore also crucial to understand the formation of electron hole puddles for both charged impurities [31, 98] and random strain fluctuations [18] as dominant mechanisms. The static dielectric function of graphene in the $T \to 0$ limit is given by [99–101]

$$\epsilon_{\mathrm{RPA}}(q = |\mathbf{k} - \mathbf{k}'|) = 1 + V_c(q)\Pi(q) \approx \begin{cases} 1 + \frac{4k_F r_s}{q} & q < 2k_F \\ 1 + \frac{\pi r_s}{2} & q > 2k_F \end{cases} \tag{2.25}$$

Here, $V_c(q) = 2\pi e^2/\kappa q$ is the Fourier transform of the bare Coulomb potential in an environment with an effective dielectric constant κ. The parameter determining the relative strength of the Coulomb interactions r_s has been introduced in Equation 2.10. $\Pi(q)$ is the static polarizability containing both interband and intraband transitions. The exact (complex) form of the dielectric function was calculated in Refs. 100, 101. The above straightforward approximations were found to be indistinguishable from the exact form [7, 98, 99]. Since graphene is gapless and its carriers behave

in a chiral relativistic manner, its screening behavior is very different compared to electrons in a conventional 2DEG [3]: $\Pi(q)$ is constant and equal to $\mathrm{DOS}(E_F)$ for both systems when $q < 2k_F$. However, for large momentum transfers with $q > 2k_F$, $\Pi(q)$ increases linearly with q due to interband transitions for graphene, whereas it quickly decreases for conventional 2DEGs [3, 100]. The linear increase enhances the effective dielectric constant for $q \to \infty$ to $\kappa\,(1 + \pi\, r_s/2)$, revealing insulator-like screening behavior at larger wave-vectors. Hence, screening in graphene is an odd combination of metal-like screening (at distances longer than λ_F) and insulator-like screening (at distances shorter than λ_F) [7, 100]. As a result, long range impurity potentials are not screened out properly in graphene. The polarizability for finite temperature was studied in Refs. 3, 102 and results in a temperature dependent static dielectric function $\epsilon(q, T)$. This gives rise to a temperature dependence of the conductivity described by a universal disorder model (see Sections 2.3.2 and 2.3.3).

The screening in graphene depends on the (average) carrier density through k_F in Eq. 2.25. Recently, the density dependence of the RPA-screening model for graphene's chiral Dirac fermions was proven by direct STM measurements of the screening length scale of a *single* charged impurity on a clean graphene surface [103]. The screening length was found to be unusually long, on the order of 10 nm. It also decreased with increasing average carrier density. For the sake of gaining an improved understanding, it can be shown in the qualitative Thomas-Fermi screening model that the Thomas-Fermi screening wave vector scales as $q_{TF} \propto n^{0.5}$ as a result of graphene's linear energy dispersion. Hence, the characteristic length scale q_{TF}^{-1} diverges at the CNP for $T \to 0$ in the disorder-free graphene case [98, 103]. The functional dependence of the bare and screened Coulomb interaction on the carrier density also remains the same apart from a multiplicative constant [3, 98]. This is in agreement with the above conclusion (for RPA) that long range disorder cannot be screened out entirely.

These findings allow us to understand and quantitatively describe the statistical properties of the density landscape composed of disorder induced electron and hole puddles at the CNP (in the $T \to 0$ limit). The ground state for interacting chiral massless carriers in the presence of screened random disorder can be treated either in a self-consistent approximation [18, 31, 98, 99] or can be obtained by brute force density functional (DFT) or Thomas-Fermi-Dirac (TFD) theory calculations minimizing the energy functional [105–108]. Since both methods are in reasonable semi-quantitative agreement[12] at the CNP [7, 29], we focus here on the self-consistent

12 The value of $n_{rms}^{SCA} = \sqrt{3} n_*^{SCA}$ calculated from the self-consistent approximation can be compared to numerical energy functional minimization calculations within TFD theory for charged

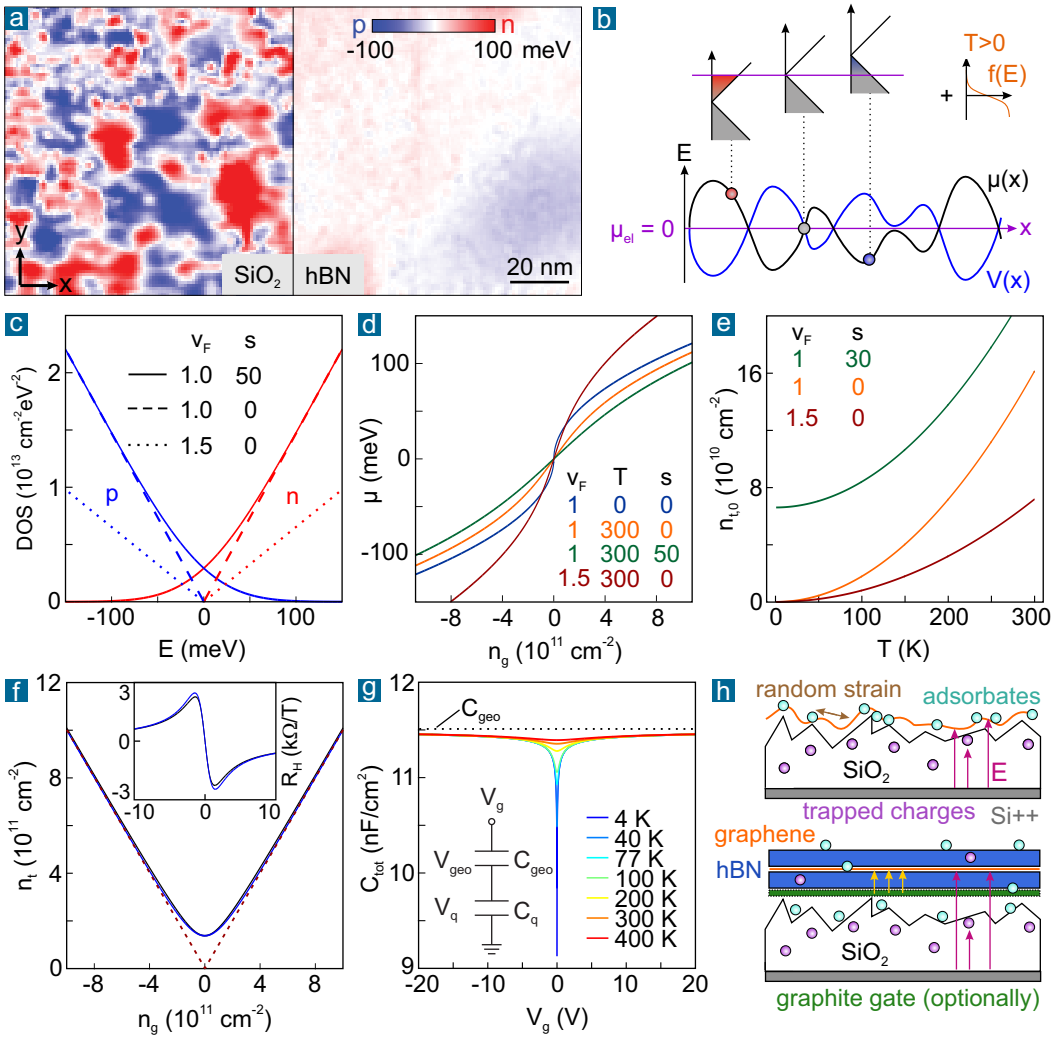

Figure 2.5. Charge carriers in graphene. a: STM map revealing the local potential in graphene at the CNP on SiO_2 and hBN at 4.5 K (modified from Ref. 104). b: Fluctuating disorder potential V(x) raises and lowers locally the Dirac point. Increasing temperature (orange) smooths the disorder induced density fluctuations. c: DOS according to Eq. 2.16 and 2.33 for different values of the Fermi velocity v_F (in 10^6 m/s) and the disorder strength parameter $s = V_{rms}$ (in meV, see Eq. 2.33). d: Average chemical potential at finite temperature T (in K) and for various sets of v_F and s. For a given gate induced average density n_g, it is calculated numerically using Eq. 2.23. e: Total carrier density at the CNP using Eq. 2.24 and 2.35. f: Total density and Hall coefficient (inset) using the channel (black) and the DOS tail model (blue line) for s = 30 meV, T = 200 K. Red line for s = 0 meV, T = 0 K. g: Quantum capacitance effect (see Section 2.2.6). C_{geo} for 300 nm SiO_2. h: Microscopic disorder mechanisms (see Section 2.2.5).

approximation (SCA) approach. The central idea behind both methods is to balance self-consistently the local doping caused by the random disorder potential and the enhancement of screening in doped graphene [7]. Accordingly, the spatially varying disorder potential induces both varying carrier density as well as local variations in the screening strength.

In SCA, the target is to find a self-consistent carrier density n_*^{SCA} for an assumed homogeneous system (with screening captured by a global RPA dielectric function) that adequately matches the observable properties of the spatially inhomogeneous real system at the CNP in the $T \to 0$ limit, such as for instance the minimum in the conductivity and its broadening (see Section 2.3.3) [7, 29, 99]. The correlation length and the amplitude of the puddles at the CNP can then be derived from the spatial auto-correlation function of the self-consistent RPA-screened disorder potential $\langle V_{dis}(r)V_{dis}(0)\rangle$. Using the self-consistent model for charged impurities, Adam et al. were able to capture experimental observations from direct probe techniques performed at low temperature [31]: *At the CNP*, the size of the puddles increases while their amplitude decreases when the impurity density is lowered, e.g. by placing graphene on hBN (as shown in Figure 2.5a). The whole system becomes smoother with large areas of almost zero density, in agreement with findings employing TFD [105]. Moreover, the self-consistent model can also explain recent STS measurements that indicate a decrease in puddle size and amplitude when the gate controlled average density n_g is tuned *away from the CNP*. A good quantitative agreement between SCA and the experiment is achieved by using the approximation $k_F = \sqrt{\pi(n_g + n_*^{SCA})}$ in the RPA dielectric function [98].

Connection of SCA to effective medium theory and to experimental observables, and definition of the disorder temperature scale

The self-consistent ansatz to determine the disorder induced residual carrier density n_*^{SCA} at the CNP in the $T \to 0$ limit in general reads [7, 18, 29]

$$\left\langle V_{dis}^2[n_*^{SCA}]\right\rangle = \left\langle E_F^2\right\rangle = \pi\hbar^2 v_F^2 n_*^{SCA}\,, \tag{2.26}$$

impurities in Ref. 105, or within Kohn-Sham-Dirac theory for corrugations in Ref. 108. The latter return directly the density profile $n(\mathbf{r})$ and all statistical quantities, e.g. the density-density correlation function and so $n_{rms}^{TFD} = \sqrt{\delta n(0)^2}$, can be derived [7, 29, 105, 106]. In a comparison done for charged impurities, n_{rms}^{SCA} and n_{rms}^{TFD} agree fairly well at (and close to) the CNP. This is especially true for low impurity density [7, 29]. The numerical TFD theory, however, is more accurate in the crossover regime towards high density and takes into account both nonlinear screening as well as many-body interactions [106].

where $V_{dis}[n_*^{SCA}]$ is the disorder potential screened by the assumed homogeneous system with carrier density n_*^{SCA}. The explicit expressions to calculate n_*^{SCA} for charged impurities and random strain fluctuations are given in Refs. 99 and 18, respectively.

One can further assume the disorder potential has a Gaussian distribution, then the root mean square of the density fluctuations (the width of the Gaussian distribution) n_{rms} at the CNP is related to the SCA residual density according to [7, 29, 80, 106]

$$n_{rms} = \sqrt{3}\, n_*^{SCA} \quad \text{(Gaussian)}. \tag{2.27}$$

We note that this value is directly connected to the root mean square of the Gaussian disorder potential distribution V_{rms} (see Section 2.2.4) [31, 97]. Hence, n_{rms} is an important parameter in the effective medium theory (EMT) which describes in general electrical transport through a spatially inhomogeneous system, such as for graphene near the CNP. It offers a better quantitative agreement to transport measurements than models that assume the system to be homogeneous (see Section 2.3.3 and Refs. 29, 80, 106, 109, 110). Graphene can be considered as inhomogeneous at the CNP, and the EMT model can be applied, as long as the disorder induced density fluctuations exceed the thermally activated carrier density. Since the density fluctuations are characterized by n_{rms}, one can define the disorder temperature scale T_* as the temperature when the thermal energy equals the Fermi energy related to n_{rms} [7, 80, 109]. Using Eq. 2.20, it is given by

$$T_* = \frac{E_F(n_{rms})}{k_B} = \frac{\hbar v_F \sqrt{\pi n_{rms}}}{k_B}\,. \tag{2.28}$$

It further has been shown that the temperature dependence of the conductivity near the CNP can be understood within an EMT model considering only charged impurities as dominant mechanism for both puddle formation and carrier scattering [80, 109]. For instance, the minimum conductivity at the CNP follows a universal scaling function up to the critical temperature T_c when thermally activated carriers outnumber the puddle carriers induced by charged impurities. The critical temperature is directly related to T_* according to [80]

$$T_c = \frac{\sqrt{3}}{\pi}\, T_* \approx 0.55\, T_*\,. \tag{2.29}$$

This critical temperature defines the transition from a spatially inhomogeneous, disorder-dominated system (captured by the SCA assuming a homogeneous system) to a charge-neutral homogeneous system with thermally activated electron and holes that coexist locally (at the charge neutrality point). For bilayer graphene with an

$n_{t,0}(T \to 0)$	E_F (meV)	T_* (K)	T_c (K)	T_*^{therm} (K)
50	83 / 18	1261 / 365	695 / 263	528 / 152
10	37 / 3.6	564 / 73	311 / 53	236 / 30
5	26 / 1.8	399 / 36	220 / 26.3	167 / 15.2
1	11.7 / 0.4	178 / 7.3	98 / 5.3	75 / 3
0.5	8.3 / 0.2	126 / 3.6	70 / 2.6	53 / 1.5
0.1	3.7 / 0.04	56 / 0.7	31 / 0.5	24 / 0.3

Table 2.1. Energy and disorder temperature scales for different residual total carrier densities at the CNP in the $T \to 0$ limit (in units of $10^{10}\,cm^{-2}$). Note, that $n_*^{SCA} \approx n_{t,0}(T \to 0)$. Values for monolayer (left values in each column) and bilayer graphene (right values) are given. $E_F(n_{t,0})$ calculations use $v_F = 10^6\,m/s$ for monolayer and $m^* = 0.033 m_e$ for bilayer graphene. T_* and T_c are calculated according to Eq. 2.28 and 2.29, respectively. In the last column an alternative definition of the disorder temperature scale is employed. It corresponds to the temperature T_*^{therm} when the thermally activated total density at the CNP of a disorder-free system (given by Eq. 2.24) is equal to the $n_{t,0}(T \to 0)$ values which are determined by disorder. For bilayer graphene we assume $E_F = \frac{\hbar^2 \pi n}{2m^*} = D_0 n$ and $n_{t,0}(T) = 2\ln(2) D_0 k_B T$ (as given in Refs. 97, 109, 111).

approximate parabolic dispersion valid for low densities, the DOS is constant and $E_F \propto n$. Accordingly, the same density fluctuations cause much smaller fluctuations in the chemical potential and the disorder temperature scale is lower (see Table 2.1).

As aforementioned, n_*^{SCA} is the disorder induced residual carrier density at the CNP in the $T \to 0$ limit for an assumed homogeneous system. In Equation 2.24 of Section 2.2.2 we had already introduced the total carrier density at the CNP $n_{t,0}(T)$ caused by thermal excitation of carriers in disorder-free, homogeneous graphene. In the following Section 2.2.4 we show that the electrical properties can be captured by phenomenological models which assume the system to be homogeneous and consider a finite residual total density at the CNP in the $T \to 0$ limit. As a consequence, the theoretical value n_*^{SCA} and $n_{t,0}(T \to 0)$, which can be obtained (estimated) from the experimental data, should be equal. This allows us to get a good estimate of the disorder temperature and energy scale in the sample through rather simple transport measurements by using $n_{rms} \approx \sqrt{3}\, n_{t,0}(T \to 0)$. Typical values for monolayer and bilayer graphene are listed in Table 2.1.

2.2.4. Phenomenological models for coexisting electrons and holes

As shown in the previous subsections, electrons and holes always coexist near the CNP in graphene devices. In the absence of disorder, the system is homogeneous and we can describe it using (spatially independent) total electron and hole densities (n_e, n_h) which are a function of temperature and the gate controlled average carrier density n_g (see Section 2.2.2). The total carrier density at the CNP $n_{t,0}$ then approaches zero for $T \to 0$. In real devices with disorder, the description based on total electron and hole densities is strictly only applicable when the temperature is well above the disorder temperature scale, that means when the system can be considered as spatially homogeneous at the CNP. At lower temperatures, the disorder induced carriers aggregate in electron and hole puddles. As we will discuss in Section 2.3.3, the boundaries between these electron and hole puddles constitute smooth, highly transparent pn-junctions. Hence, the electrical transport in the inhomogeneous system can be understood to good approximation by an assumed homogeneous system described by total electron and hole densities. These total densities (n_e, n_h) then substitute the spatially inhomogeneous density profile as shown in the following.

Ad-hoc residual total carrier density at the CNP and the channel approximation model

A straightforward phenomenological approach to model the inhomogeneous system by a homogeneous one is to introduce an ad-hoc residual total carrier density at the CNP $n_{t,0}$ which is finite at all temperatures since both disorder induced carriers and thermally activated carriers contribute. The average carrier density $n = n_e - n_h = n_g = \beta\,(V_g - V_{CNP})$ is fixed by the gate voltage. Then, the remaining question is how one can model the total carrier density $n_t = n_e + n_h$ as a function of n_g (or gate voltage). By the above definition, it equals $n_{t,0}$ at the CNP (for $n_g = 0$), and will approach the gate controlled average carrier density at high densities in the single carrier type regime (when $E_F \gg k_B T, k_B T_*$). In the literature, e.g. in Refs. 112–114, it was suggested to approximate the total carrier density in the graphene channel by

$$n_t = n_e + n_h = \sqrt{n_{t,0}^2 + n_g^2} = \sqrt{n_{t,0}^2 + (\beta(V_g - V_{CNP}))^2}\,. \tag{2.30}$$

This channel approximation model and $n_g = n_e - n_h$ can then be rewritten in order to obtain separate expressions for the (total) electron and hole densities in the assumed homogeneous system

$$n_e = 0.5\,(\sqrt{n_{t,0}^2 + n_g^2} + n_g) \quad \text{and} \quad n_h = 0.5\,(\sqrt{n_{t,0}^2 + n_g^2} - n_g)\,. \tag{2.31}$$

As we will demonstrate in Section 2.3.3, these densities can be used to calculate the longitudinal resistivity and the Hall coefficient using the two carrier type Drude model (which assumes the system to be homogeneous). The residual total carrier density at the CNP $n_{t,0}$ then serves as fitting parameter to the experimental data. Due to the thermal activation of carriers, $n_{t,0}$ is expected to increase quadratically with temperature. In our picture of a homogeneous system, it can be modeled by adding a constant $n_{t,0}(T \to 0)$ to Equation 2.24 describing the thermal excitation of carriers. This finite low temperature limit is determined by the disorder and constitutes a measure for its strength, or in other words, for the sample quality (for our best hBN-encapsulated device we typically obtain $n_{t,0}(T \to 0) \approx 1 \cdot 10^{10}\,\mathrm{cm}^{-2}$). The formula for the residual total carrier density at the CNP then reads

$$n_{t,0}(T) = n_{t,0}(T \to 0) + \frac{\pi k_B^2}{3\hbar^2 v_F^2} T^2 \,. \tag{2.32}$$

The low temperature offset $n_{t,0}(T \to 0)$ is either determined by fits to low temperature transport data or it can be interpreted as the SCA density n_*^{SCA} and calculated for the dominant microscopic disorder mechanism (as mentioned in the previous subsection). Moreover, the disorder temperature scale can be estimated (using $n_{rms} \approx \sqrt{3}\, n_{t,0}(T \to 0)$ in Eq. 2.28 and 2.29).

Modified DOS model to account for disorder induced inhomogeneity

So far, an ad-hoc total carrier density at the CNP was introduced in a homogeneous model system in order to account for the spatial inhomogeneous landscape of electron hole puddles in real devices. Li and coworkers (Refs. 97, 115) proposed a more elaborate model which accounts for the disorder induced inhomogeneity by a modified DOS. The latter then replaces the regular DOS in the calculations of the (spatially independent) total electron and hole densities (n_e, n_h) at finite temperature for the assumed homogeneous system. The random fluctuations in the local electrostatic potential $V(\mathbf{r})$ are assumed to obey a Gaussian distribution P(V) with standard deviation $s = V_{rms}$. The authors conclude that modifying the regular DOS (Eq. 2.16) with a Gaussian tail is an excellent approximation to the self-consistent DOS in the presence of disorder [115]. For electrons the modified DOS reads

$$D_e(E) = \int_{-\infty}^{E} D_0\,(E - V) P(V) dV = D_0 \left[\frac{E}{2} \mathrm{erfc}\left(-\frac{E}{\sqrt{2}s} \right) + \frac{s}{\sqrt{2\pi}} \exp\left(-\frac{E^2}{2s^2} \right) \right] . \tag{2.33}$$

The modified DOS for holes is calculated in a similar fashion by replacing E with -E and the resulting modified DOS with the Gaussian tails is shown in Figure

2.5c. Note, the modification of the DOS is phenomenologically motivated and not sensitive to any specific microscopic disorder mechanism. In analogy to Eq. 2.22, the expression for the (total) electron density is then given by

$$n_e = \int_{-\infty}^{\infty} D_e(E) \frac{dE}{e^{\beta(E-\mu)} + 1}, \tag{2.34}$$

and correspondingly for the hole density. Again in analogy to Section 2.2.2, the chemical potential μ is calculated considering charge carrier conservation with the condition $n_e - n_h \stackrel{!}{=} n_g$. Panel d of Figure 2.5 shows the dependence of the chemical potential on the gate controlled density for a given disorder parameter s. This condition also justifies the introduction of the tails in the DOS which at first glance appear to be artificial and physically wrong (since additional states are introduced). However, since μ adapts accordingly, the electron DOS tail can be interpreted as anti-hole states and vice versa. All in all, the electron and hole densities of the assumed homogeneous system, which is used to model the properties of the electron hole puddles in real devices, are a function of gate controlled average density, temperature, and disorder parameter s. The latter then constitutes a measure of the disorder strength (higher quality devices are characterized by smaller values).

The introduction of the ad-hoc residual density and the channel approximation give us analytical expressions for the dependence of n_e and n_h on the gate controlled density (Eq. 2.31). In contrast, the DOS tail model typically requires numerical computations of the integrals. For given temperature and disorder parameter, the chemical potential first has to be determined numerically for each gate controlled average density. Then, the electron and hole densities can be calculated. However, exactly at the CNP, analytical expressions for n_e and n_h can be derived and the total carrier density at the CNP in the DOS tail model is given by [97]

$$n_{t,0}^{\mathrm{DOStails}} = \frac{s^2}{\pi\hbar^2 v_F^2} + \frac{\pi k_B^2}{3\hbar^2 v_F^2} T^2 . \tag{2.35}$$

By comparing this formula to Eq. 2.32, both models will obviously result in the same total carrier density at the CNP, when $n_{t,0}(T \to 0) = s^2 (\pi\hbar^2 v_F^2)^{-1}$. For example, $s = 30\,\mathrm{meV}$ corresponds to an ad-hoc $T \to 0$ density of $6.6 \cdot 10^{10}\,\mathrm{cm}^{-2}$ (when $v_F = 10^6\,\mathrm{m/s}$). Moreover, both models will converge at high densities (towards the gate controlled density in the single carrier type regime). The difference between the two models is also negligible in the crossover regime (for intermediate densities), when the disorder parameter s and $n_{t,0}(T \to 0)$ are chosen accordingly (as demonstrated in panel f of Figure 2.5). The total carrier density in the channel model is slightly larger at intermediate densities which results in slightly smaller Hall coefficients (calculated using the two carrier type Drude model equations given

in Section 2.3.3). Since both the disorder parameter s and $n_{t,0}(T \to 0)$ are anyways used as fitting parameters, this minor deviation is negligible. However, the analytical forms, given by the channel approximation, allow a much faster fitting of the measured magneto-transport curves. Hence, we will employ the channel approximation with the ad-hoc residual total carrier density at the CNP in this work (see Chapter 4).

2.2.5. Microscopic origin of the extrinsic disorder

In the previous subsections, the theoretical framework of the electron hole puddle formation caused by extrinsic disorder has been discussed in general (not restricted to a certain microscopic disorder mechanism). Moreover, we have presented phenomenological approaches which describe the spatially inhomogeneous landscape by means of homogeneous model systems with total electron and hole carrier densities. These model systems then exhibit a residual carrier density at the CNP (or an equivalent parameter) typically chosen such that the measured transport properties of the disordered graphene are reproduced[13]. In the following we address the important question about the microscopic origin of the extrinsic disorder. This information is required to go beyond phenomenological models. Potential microscopic disorder models are then implemented in the above system description, such as the self-consistant approximation, the effective medium theory, and the density functional theory, in order to obtain quantitative predictions which can be compared to experiments. Moreover, if the extrinsic cause of the disorder is understood, manufacturing and processing steps as well as the device design, such as the choice of substrate or encapsulation, can be optimized to improve sample quality and achieve extremely low extrinsic disorder (see Chapter 3).

So far, we have only considered a general concept of extrinsic disorder as the cause for the formation of electron hole puddles near the CNP. This spatial inhomogeneous landscape can be examined directly using local probe techniques. The amplitude of the density fluctuations and other statistical parameters, for example the puddle correlation length, are obtained by scanning the sample. These quantities can then be compared to the theoretical predictions assuming a certain microscopic disorder mechanism [31, 98]. Obviously, those experiments should be conducted at low temperature, such that thermal generation of carriers does not smear out the disorder

13 These concepts are revisited in Section 2.3.3 where we will introduce the two carrier type Drude model, which assumes a homogeneous system, in order to describe electrical transport in graphene. Moreover, the effective medium theory will be discussed which explicitly accounts for the spatial inhomogeneity of the system near the CNP.

induced density fluctuations. At low temperature, the extrinsic disorder also constitutes the dominant source of charge carrier scattering for a wide density range starting from the CNP (when phonon scattering can be neglected). The density dependence of the conductivity then offers another testbed for potential microscopic theories. We will discuss the diffusive transport regime and various scattering mechanisms in the subsequent Section 2.3. The main experimental finding is that the conductivity shows a linear dependence with density before it saturates near the CNP due to the disorder induced residual carrier density. This implies that the mobility does not depend on the density. This finding and the formation of electron hole puddles, together, can be explained only by a long range disorder potential[14]. A third type of experiment, the measurement of quantum interference effects, provides further information about the dominant microscopic source of disorder. As discussed in Appendix B.1, quantum interference effects allow to extract several scattering times by fitting the magnetoresistance curves. Intervalley, symmetry conserving intravalley, and symmetry breaking intravalley scattering involving a pseudospin flip can be distinguished (see also Section 2.1.5). Recent experimental findings suggest that symmetry breaking intravalley scattering prevails at low temperature (see Appendix B.1 for details).

In summary, a candidate microscopic theory needs to explain the experimentally observed electron hole puddle landscape, the linear density dependence of the conductivity, and the quantum interference effects at low temperature. In order to allow the identification of the underlying microscopic disorder mechanism, graphene has been studied in different environments, for instance by placing it on different substrates or by encapsulating it in between different materials. These substrates are characterized by their intrinsic charged impurity density n_{imp}, their dielectric constant and their surface roughness. Based on the analysis of a multitude of experiments, discussed in Section 2.3.2, in Chapter 3, as well as below, one can conclude that charged impurities [29], in close proximity as well as larger distance from graphene, and random strain fluctuations [18] are the two most likely sources of microscopic disorder. However, none of these two sources of disorder alone can explain all reported findings. Unfortunately, these two remaining candidates are often difficult to disentangle in real devices. A comprehensive study, performing all three key experiments (scanning local probe measurements addressing the characterization of the electron hole puddles, transport and quantum interference) together and systematically on various substrates but for the same processing conditions, is lacking. Still, the focus on these two microscopic disorder mechanisms is widely supported by the published results. The next paragraph summarizes insights that have been gained

14 See also the comment on short range disorder potentials in the footnote at the beginning of Section 2.2.3.

by local probe measurements on the electron hole puddle formation. The discussion about the dominant disorder mechanism will be continued in Section 2.3.2 where charge carrier scattering is addressed.

Insights from scanning local probe techniques and charged impurity disorder

Local probe techniques provide direct access to the electron hole puddle landscape and have been the unambiguous experimental proof for its existence [96]. Different techniques, such as SET [96], STM/STS [98, 104, 116–120], and AFM-Kelvin probe microscopy [121], were employed to study the density fluctuations in graphene on SiO_2 and hBN as well as the intrinsic charged impurity concentration in those substrates. Typically, such techniques require an uncovered sample surface in order to achieve a sufficient (or even atomic scale) spatial resolution[15]. Unfortunately, the sample surface can then become contaminated during processing steps (while this is not the case for encapsulated graphene). Most of these residues and adsorbates, however, can be removed in-situ by annealing in vacuum prior to conducting scanning probe measurements. The experiments are carried out at low temperature such that thermal charge carrier generation is negligible.

Scanning probe measurements typically find density fluctuations near the CNP with a magnitude of a few $10^{11}\,\mathrm{cm}^{-2}$ for graphene on SiO_2, and one to two orders of magnitude smaller for graphene on hBN. These values are similar to the residual total carrier densities at the CNP obtained from transport measurements (see Chapter 4). To our knowledge, no local probe experiments have been conducted on very fragile (current annealed) suspended devices which show the lowest reported residual carrier density in transport experiments (below $10^{9}\,\mathrm{cm}^{-2}$ [126]). STM experiments

15 In particular, STM/STS measurements rely on direct access to the surface but then offer atomic resolution. SET measurements can also be deployed for samples equipped with a thin encapsulation layer (e.g. only few atomic layers of hBN), since its spatial resolution is already limited by the size of the SET. Unconventional optical techniques might overcome this limitation and allow studying devices with thicker encapsulation layers which are normally used in devices intended for electrical transport measurements. Recently, near field photocurrent nanoscopy was proposed and employed to study charge puddles in hBN-encapsulated graphene [122]. Moreover, other measurement techniques can be employed which are not able to observe directly the electron hole puddles but provide valuable indications on the doping and strain distribution in graphene. Scanning confocal Raman spectroscopy allows probing nanometer-scale strain variations as well as doping fluctuations (which are, however, averaged over the laser spot size) [123, 124]. The roughness of graphene on a certain substrate or of the substrate surface itself can be determined by AFM measurements (see Chapter 3). TEM diffraction patterns reveal information on the microscopic corrugations in freestanding graphene [16] and on the interface roughness of (hBN-)encapsulated graphene [125].

were also able to identify a spatial correlation between the electron hole puddles in graphene and charged impurities located near the surface of the SiO_2 substrate [118]. Kelvin probe measurements, performed on bare substrates (without graphene), state a one to two orders of magnitude lower intrinsic charged impurity density in exfoliated hBN layers than in thermally grown SiO_2 [121]. This agrees very well to the observed decrease of the density fluctuations in graphene on hBN by the same order of magnitude [104, 120]. These findings clearly point to charged impurities trapped near the substrate surface as microscopic disorder mechanism. As mentioned in Section 2.2.3, charged impurities have been considered for both the SCA model (using RPA theory for screening) as well as for DFT (or TFD) calculations. They nicely explain that, for graphene at the CNP, the amplitude of the electron hole puddles decreases while their lateral dimension increases when the charged impurity density n_{imp} of the substrate becomes smaller [31]. Panel a of Figure 2.5 shows two exemplary STM maps of graphene on SiO_2 (high n_{imp}) and hBN (low n_{imp}) from the literature (Ref. 104). The average size of the puddles at the CNP is on the order of ten nanometer on SiO_2 while they are substantially larger on hBN. Moreover, the transition between the puddles is much smoother on hBN revealing large areas with almost zero carrier density (white color in the images). Hence, a significant part of the system approaches the intrinsic Dirac cone physics in the limit of very small energies. Finally, the SCA model considering charged impurity disorder also explains quantitatively the observed decrease in amplitude and size of the puddles when the average density is tuned away from the CNP [98].

The microscopic origin of charged impurities is also sketched in panel h of Figure 2.5. As aforementioned, charged impurities can be located near the substrate surface (or, more general, near the interface between graphene and the encapsulating material). Moreover, physical adsorbates attached to graphene's surface or trapped between the interfaces, such as residues from device processing or airborne contaminants, constitute other plausible sources of charged impurities. We note that charged impurities not only induce the electron hole puddles but also a net charge transfer may occur which is reflected in an initial doping of graphene (leading to a non-zero V_{CNP}, as discussed in Section 2.2.1 and Chapter 3.2). The microscopic theory of charged impurities is then able to explain self-consistently this initial doping [29, 99]. Finally, transport measurements reveal that the already good quality of hBN-encapsulated graphene devices can be improved further, in terms of the residual total density at the CNP and the mobility, when a graphitic gate (i.e. a multilayer graphene flake) instead of a silicon back gate is used (see Chapter 3 for details). This finding likely implies that remote charged impurities also play a role in such clean devices with low charged impurity density in vicinity to graphene (near the hBN-graphene interface). The remote charged impurities are located at the SiO_2-hBN interface and, hence, have a detrimental impact on the field effect

homogeneity when the silicon back gate is used (yellow and magenta electric field arrows in Figure 2.5h). When employing a graphitic gate (i.e. gating without the SiO_2-hBN interface), these remote charged impurities are then screened. Unfortunately, to our best knowledge, no scanning probe experiments have been conducted on a single device which is partially gated by the conventional silicon back gate and partially by a graphitic gate.

Insights from scanning local probe techniques and random strain fluctuations

An alternative (or complementary) microscopic disorder mechanism is related to graphene's high out-of-plane flexibility. AFM studies report that graphene conforms almost completely to the corrugations on the underlying substrate (see Refs.127, 128 and our measurements in Section 3.1.3). Our AFM measurements find the surface roughness of few tens of nanometer thick hBN flakes to be at least four times smaller than for SiO_2 (ultimately, our resolution is limited by the noise level of the AFM system). STM studies even claim atomically flat graphene on hBN within the observed scan area of few tens of nanometer [104, 120]. This is supported by cross-sectional TEM images taken on the hBN-graphene interface (for areas without trapped contamination [129]) as well as by the sharpness[16] of TEM diffraction pattern spots [125].

Hence, random strain fluctuations induced by graphene's conformation to the substrate have been proposed as the prevailing source of microscopic disorder [18, 108, 117, 131, 132]. The reduction of the amplitude of the electron hole puddles as well as their spatial broadening on hBN in comparison to SiO_2 have been (qualitatively) attributed to the smaller roughness of the substrate surface leading to smaller strain fluctuations [18, 108]. We have also illustrated this idea in panel h of Figure 2.5.

16 When the substrate below graphene is removed and graphene is freestanding, diffraction pattern spots broaden and reveal intrinsic ripples with out-of-plane deformations reaching an amplitude of up to one nanometer [16]. This intrinsic roughness is attributed to thermal fluctuations at room temperature which also limit the carrier transport through flexural phonon scattering in these suspended devices [17, 130] (see also Section 2.3.2 and Chapter 3). However, at low temperature, these fluctuations should freeze out and both the mobility as well as the residual carrier density at the CNP exhibit record values, as reported in transport studies [126]. Unfortunately, owing to challenging technical difficulties, no experiments were carried out so far attempting to measure the intrinsic roughness and the electron hole puddle landscape of suspended graphene at low temperature (in the absence of thermal fluctuations). These efforts would "only" help to prove the microscopic disorder mechanism but, in commercial applications, suspended graphene will likely never be relevant. Moreover, at room temperature, the mobility of freestanding graphene is low due to flexural phonon scattering [130]. Accordingly, we focus here on supported graphene.

The random strain fluctuations picture is also supported by scanning confocal Raman spectroscopy [123, 124]. As we will discuss in detail in Section 3.2.1, the Raman spectra of graphene reveal information on the nanometer-scale strain variations (averaged over the laser spot size). In agreement with AFM measurements, they are found to be small on substrates such as hBN on which graphene exhibits low residual carrier density at the CNP and high mobility in transport. In the most recent[17] theory of the random strain fluctuations, both in-plane and out-of-plane deformations are considered. They generate both a scalar V_s and a vector disorder potential $\mathbf{A}$ in the Hamiltonian [18, 131]. As discussed in more detail in Section 2.3.2, the formation of the electron hole puddles is related to the scalar potential (with self-screening described by RPA), while the scattering of the charge carriers is mainly governed by the vector potential. The vector potential has long range character and causes a linear density dependence of the conductivity. More important, the vector potential is able to differentiate between the two sublattices and thus can influence the pseudospin, i.e. it constitutes an intravalley symmetry breaking scattering mechanism[18]. This agrees well to the previously summarized findings of quantum interference experiments which imply a dominant symmetry breaking disorder mechanism. On the contrary, these findings cannot be explained by the charged impurity disorder model which allows only for symmetry conserving scattering [18][19].

Unfortunately, random strain fluctuations have not yet been incorporated in the above mentioned general theories, such as SCA, in order to obtain quantitative (and analytical) results, describing the electron hole puddle landscape, which could be compared to scanning probe experiments. So far, only DFT-based calculations using experimentally obtained height-fluctuation maps have been performed and compared to the density profile in graphene on SiO_2 [108]. Hence, in general, the strain fluctuations picture is not as mature and widely used as the charged impurity theory. It also bares some weaknesses. For instance, it lacks to explain the initial (global) doping which one observes in many (non-encapsulated) devices. Obviously, a charge transfer to charged impurities is the reason for this doping. Since this evidences the presence of a notable amount of charged impurities in such devices, we

17 The first studies considered only out-of-plane deformations referred to as ripples (e.g. in Ref. 133). Only recently, the complete picture of random strain fluctuations has been discussed [18, 108, 131].

18 As a result, small momentum backscattering within a single valley is no longer suppressed. We wish to remind that the inherent suppression of intravalley backscattering in graphene in the absence of this vector potential is one important reason for the large room temperature mobility in graphene.

19 The authors of Ref. 18 further claim that the relation between the residual carrier density at the CNP and the mobility favors only the random strain fluctuation mechanism to model the disorder in graphene as discussed in Section 2.3.2. However, such a relation can also be justified in the charged impurity density model (even though this has not yet been shown quantitatively).

surely expect them to contribute considerably to the scattering and the formation of the electron hole puddle landscape. The device quality improvement observed in devices with a graphitic gate is also not consistent with the random strain fluctuations model since these fluctuations remain the same for sufficiently thick hBN layers, independent of the presence of an additional multilayer graphene flake (used as the graphitic gate) underneath the hBN. Instead, this points to remote charged impurities. Another weak spot of the strain fluctuations model is that, although graphene on mica is atomically flat [134], its mobility remains low [135] (STM experiments measuring density fluctuations on mica have not been conducted). Instead, the intrinsic charged impurity density in mica is found to be high [136]. This explains the low mobility indicating that random strain fluctuations are in general not the only culprit for the disorder on all substrates.

Concluding remarks

To conclude, the microscopic origin of the disorder is still partially puzzling. It is likely that not one mechanism alone can be regarded as generally dominant, although a certain hierarchy is obvious: Charged impurities definitely play a major role in (non-encapsulated) graphene on substrates containing a large charged impurity density, e.g. oxides (SiO_2, mica). As recently observed by STM measurements [137], this also holds for graphene on transition metal dichalcogenides (TMD) which are found to exhibit a considerable impurity density. Moreover, encapsulated graphene shows self-cleansing of its interfaces only on certain atomically flat substrates, e.g. hBN and TMD, but not on atomically flat oxides such as mica (see Ref. 138 and Chapter 3 for a detailed discussion of the self-cleansing effect). Accordingly, even atomically flat substrates may possess a rather high intrinsic charged impurity level (e.g. TMD) and contaminants, such as air-borne adsorbates or residues from device processing, can remain trapped between the interface of graphene and the substrate. These contaminants then constitute a source for charged impurities and also reduce significantly the flatness of graphene (even in the case of substrates showing the self-cleansing effect, the contaminants aggregate in bubbles, as observed by cross-sectional TEM [129, 138]). Hence, only for flat substrates exhibiting low intrinsic impurity density and high affinity to graphene (such that self-cleansing occurs), random strain fluctuations and, eventually, remote impurities constitute the dominant source of disorder in graphene. They then set the ultimate lowest limit of disorder level that can be achieved in real devices on a substrate. As of today, all these requirements are only met by hBN (see Section 3.1.3).

Several experiments also report that the chemical reactivity of graphene and the initial doping level are linked to the substrate roughness and its functional surface groups [95, 136, 139]. This constitutes another argument for the interplay between strain fluctuations and charged impurities: strained π-bonds of graphene stronger attract adsorbates, and adsorbates also deform locally the graphene lattice. Hence, random strain fluctuations and charged impurities are often difficult to disentangle and a common theoretical picture is lacking due to its high level of complexity. Finally, the disorder strength and so the amplitude of the density fluctuations are, to a certain extent, also subject to statistical variations within an equally processed batch of devices (see Chapter 3). This makes systematic studies not easier and reliable conclusions have to be made probing a large ensemble of devices.

2.2.6. Influence of the quantum capacitance

In the previous sections we have neglected quantum capacitance effects when studying the carrier density tuning by a gate voltage. Assuming the metallic gate electrode has a much larger DOS compared to the graphene layer, the system's total capacitance can be described by the geometrical and the quantum capacitance of graphene connected in series, as shown in Figure 2.5g. In the disorder-free case at low temperature, the quantum capacitance is proportional to the DOS which due to graphene's linear band structure is not constant (in contrast to what occurs in conventional 2DEGs and bilayer graphene) but depends on the Fermi energy controlled by the gate voltage. Hence, the quantum capacitance in monolayer graphene can become very small around the CNP and will become the main contribution to the system's total capacitance when $C_q \leq C_{geo} \approx 11.5\,\mathrm{nF/cm^2}$ (for 300 nm SiO_2). However, any residual carrier density at the CNP, either induced by disorder or thermally generated, will lead to a finite quantum capacitance, even when $\mu \to 0$. At finite temperature but in the absence of disorder, the dependence of the quantum capacitance on the gate voltage is given by the following self-consistent expression [140–142]

$$C_q = \frac{2e^2 k_B T}{\pi(\hbar v_F)^2} \ln\left(2 + 2\cosh\frac{eV_q}{k_B T}\right) \quad \text{and} \quad V_q = V_g\left(1 - \frac{C_q}{C_{geo} + C_q}\right). \tag{2.36}$$

Figure 2.5g depicts the gate voltage dependence of the total capacitance where we have assumed the typical equivalent oxide thickness (EOT) of the devices studied in this work. Only for very small temperatures and gate voltages of a few hundred mV, the total capacitance deviates notably from the geometrical capacitance. When one in addition accounts for disorder with equivalent temperature scale on the order of 100K (reached in our best devices, see also Table 2.1), the quantum capacitance in monolayer graphene remains around one μF at the CNP, as shown theoretically [143]

and observed in quantum capacitance measurements [144–146]. This is at least two orders of magnitude larger than C_{geo} for our typical EOT, and allows us to neglect the influence of quantum capacitance in our studies. Note, in bilayer graphene C_q is constant and also much larger than typical values of C_{geo} [82].

However, quantum capacitance effects will play a significant role close to the CNP in devices with low disorder and/or small EOT at low temperature, e.g. for ionic gating [147] or for devices with graphitic or metal gates using only thin hBN flakes as dielectric. The latter configuration can be used to intentionally conduct capacitance spectroscopy measurements as performed to study the behavior of graphene/hBN superlattices [148] and many body interaction phenomena [71].

2.3. Diffusive transport

The electrical (magneto-)transport in graphene, one of the main topics in this work, can be mostly described by semi-classical theory in the diffusive regime (for non-quantizing magnetic fields). Quantum interference corrections to the Drude-Boltzmann conductivity are found to be negligible ($\approx 1\%$), but they can provide valuable indications on the dominant scattering mechanism, as discussed in Appendix B.1 (which also gives an overview of the quantum Hall effects in graphene emerging for quantizing magnetic fields). Only for clean devices with extremely low extrinsic disorder many-body interactions become relevant at intermediate temperatures (typically between 50 and 200 K) and manifest in collective transport phenomena. The occurrence of ballistic transport at low temperature depends on the device dimensions and also requires a low disorder level such that the mean free path L_{mfp} exceeds the device dimensions (at higher carrier densities). These two mesoscopic transport regimes are described in Section 2.4 in more detail. The main conclusion is that the electrical transport at low densities near the CNP, the density range we mainly focus on in this thesis, remains diffusive[20] over the whole temperature range for all real graphene devices (see also Section 2.5).

The semi-classical Drude-Boltzmann theory, employed to describe diffusive transport in graphene, is strictly speaking only applicable when $k_F L_{mfp} \gg 1$ [7]. This condition is clearly fulfilled only at higher densities. The main question is then what will happen when the carrier density approaches the Dirac point (the charge neutrality point). Early theoretical studies considering disorder-free[21] graphene predicted a quantum-ballistic universal conductivity or an Anderson localization induced insulator transition at the CNP in the $T \to 0$ limit [3, 7]. However, as discussed in the previous section, a fundamental consequence of the unavoidable (long-range)

20 At low temperature, below the disorder equivalent temperature scale, both the inhomogeneous landscape and the carrier scattering is governed by the extrinsic disorder near the CNP, leading to diffusive transport along the puddles. For most devices, the disorder level is so high that it remains the dominant scattering mechanism even beyond room temperature and then, at higher temperatures, is replaced by phonon scattering. Above the disorder temperature, the system at the CNP is then homogeneous (with thermally activated, coexisting electrons and holes) but the carrier transport is diffusive. For ultra clean devices with low disorder, the disorder temperature scale can become so low that the system turns homogeneous before phonon scattering starts to play an important role. Then, carrier-carrier interactions determine the carrier transport at the CNP. However, even in this particular case, electron hole scattering processes cause the system to remain diffusive (see also Refs. 80, 81). Accordingly, ballistic (or viscous) effects may emerge only at higher densities, since the low density region near the CNP is always diffusive either as a result of disorder, electron-hole or phonon scattering.

21 Or, to be more exact, these theories did not consider that (random) disorder leads to the inhomogeneous electron hole puddle landscape at the CNP.

extrinsic disorder in real devices is that the system breaks up into a spatially inhomogeneous landscape consisting of electron hole puddles for temperatures below the disorder equivalent temperature scale[22]. Then, electrical transport at low density is governed by the physics of this inhomogeneous system with spatially varying density, that means by electron-rich and hole-rich metallic regions whose boundaries constitute smooth pn-junctions and are highly transparent and conducting [3, 31]. Consequently, the intrinsic Dirac point physics is obscured. Inside these puddles, the above condition for the Boltzmann theory is fulfilled to good approximation. The Boltzmann theory then serves as an important component describing the scattering processes in refined theories which account for the disorder induced spatial inhomogeneity near the CNP, such as the self-consistent approximation theory and the effective medium theory [7, 29]. The latter capture the important experimental observation that the conductivity at the CNP approaches a minimum $\sigma_{\mathrm{min}}(\mathrm{T} \to 0)$ in the low temperature limit which is non-universal and sample-dependent [7]. This can be successfully explained by the extrinsic disorder alone which both induces the residual carrier density and causes scattering.

A reduction of the disorder level then leads to a lower residual carrier density (in the low temperature limit) as well as to an increase of the mobility. Hence, $\sigma_{\mathrm{min}}(\mathrm{T} \to 0)$ at first increases when the disorder level drops from high (on SiO_2) to moderate values (on hBN), since the mobility enhancement overcompensates the decrease of residual carrier density. Ultimately, however, the loss of charge carriers prevails and devices with extremely low disorder level (e.g. hBN-encapsulated or suspended) exhibit the smallest $\sigma_{\mathrm{min}}(\mathrm{T} \to 0)$ values [149]. Moreover, for these high quality devices, the minimum conductivity and the width of its plateau (the density range around the CNP where the conductivity saturates towards σ_{min}) substantially increase with temperature due to thermal activation of charge carriers while the mobility limited by the extrinsic disorder decreases only slowly. Accordingly, cleaner devices show a much more pronounced insulating behavior at the CNP than dirtier devices[23], since the thermally activated carriers outnumber the disorder induced ones already

22 For conventional devices on SiO_2 the system remains inhomogeneous and disorder-dominated at the CNP even above room temperature. For the best encapsulated devices the disorder equivalent temperature scale is on the order of 100 K. The lowest disorder level ever reported has been achieved for suspended graphene but, even in this case, intrinsic Dirac point physics in the low temperature limit is obscured, since the disorder equivalent temperature scale is still about 10 K [126].

23 However, the minimum conductivity always saturates towards a sample-specific $\sigma_{\mathrm{min}}(\mathrm{T} \to 0)$ below a certain temperature (defined by the disorder level) and does *not* continue to decline as the temperature is lowered further. The saturation distinguishes this insulating behavior from Anderson localization or from an exponential behavior related to a band gap formation (see also Ref. 149).

at moderate temperatures. All in all, the electrical transport and its temperature dependence in this two carrier type regime near the CNP are highly complex.

At higher carrier densities, away from the CNP, only electrons or holes are present. The system in this single carrier type regime is essentially spatially homogeneous and can be described by the Drude-Boltzmann theory. Transport measurements find the conductivity to increase linearly with density until this relation eventually becomes sublinear at high density. In particular for devices with low disorder and at higher temperatures, the conductivity at high density even tends to saturate towards a maximum. At fixed density, the conductivity decreases with increasing temperature which corresponds to metallic behavior[24]. The complex density dependence of the conductivity is understood by the interplay of different scattering mechanisms. Disorder scattering dominates at low densities and, depending on the disorder strength, also up to high densities. As already discussed in the previous section, the most likely extrinsic disorder mechanisms, such as charged impurities and random strain fluctuations, have long-range character leading to a density-independent mobility (to a conductivity that depends linearly on density). The density-independent contribution to the conductivity stems from short range scatterers, for instance lattice defects, ripples, and acoustic phonons. Although electron-phonon coupling is weak in graphene, phonon scattering exceeds over disorder scattering at high enough densities and temperatures. In clean devices with low disorder level, that means with a high disorder-limited mobility (steep increase of the conductivity with density), acoustic phonon scattering then constitutes the intrinsic mobility limit at room temperature for densities on the order of $10^{12}\,\mathrm{cm}^{-2}$.

This section also covers magneto-transport phenomena in the diffusive regime (for non-quantizing magnetic fields). While the semi-classical theory predicts the absence of magneto-resistance and the observation of the trivial Hall effect ($R_H \propto 1/n$) in the single carrier type regime, the coexistence of two different types of charge carriers near the CNP leads not only to the minimum conductivity but also to a strong classical magneto-resistance and a vanishing Hall coefficient. Accordingly, at the crossover from the single carrier type to the two carrier type regime the Hall coefficient will exhibit a maximum. Assuming the system to be spatially homogeneous, the two carrier type Drude model can be employed to describe these observed phenomena. This works surprisingly well and the deviations from the effective medium theory, which accounts for the spatial inhomogeneity, are small (at least for classically weak magnetic fields when $\mu^2 B^2 \ll 1$).

24 Hence, graphene can be also considered as a system with density-tunable metal-insulator transition (with respect to the temperature dependence of the conductivity). This corresponds to the crossover between the single carrier type and two carrier type regimes, respectively.

The section is structured as follows. We start with the introduction to the semi-classical Drude-Boltzmann theory demonstrating the difference between short and long range scattering potentials. Then, all so far proposed microscopic scattering mechanisms are discussed and the most important ones are identified. Subsequently, the transport in the two carrier type regime is considered before we finally summarize the temperature dependence of the conductivity in the different regimes.

2.3.1. Semi-classical theory in the high density regime

In the following we consider graphene in the high density homogeneous regime ($E_F \gg k_B T, k_B T_*$) when only one carrier type is present.

Definitions of magneto-transport parameters in the semi-classical picture

The Drude model, which assumes the charge carriers to have a mass, can be adapted to the massless quasiparticles (with charge $q = \pm e$) in graphene by using a more general definition of the Drude mobility[25] μ and the Drude conductivity σ:

$$\mu \equiv \frac{v_D}{E} \text{ and } j = qnv_D = qn\mu E \equiv \sigma E \quad \leftrightarrow \mu = \frac{\sigma}{q\,n}\,. \tag{2.37}$$

Here, v_D is the drift velocity in the stationary case, E the external electric field, j the current density, and $n = n_g$ the (average) carrier density which is typically controlled in experiments by the gate voltage (see Section 2.2). Hall cross or Hall bar geometries are used in this work to determine the transport coefficients in graphene. Let L be the length and W be the width of the channel (as shown in Figure B.2 in Appendix B.2), then one can derive the field effect mobility μ_{FE} by measuring the resistivity $\rho = \sigma^{-1} = R\,W/L$ and determining the carrier density using the gate capacitance model (Eq. 2.18). The field effect mobility is then defined by

$$\mu_{FE} \equiv \frac{1}{e\,n_g\,\rho} = \frac{\sigma}{e\,n_g}\,. \tag{2.38}$$

It is obvious that this definition is only valid if one carrier type is available. At the CNP ($n_g \to 0$), the field effect mobility diverges, since this definition does not

25 In the standard Drude theory for charge carriers with a mass, an ad-hoc scattering time τ is introduced. For graphene, one introduces the mobility as an ad-hoc parameter (as relation between drift velocity and electric field). The Boltzmann theory allows to calculate the scattering time for an arbitrary microscopic mechanism and provides a general expression for the (Boltzmann) conductivity, not restricted to carriers having a mass. The mobility is then derived from the general relation $\mu = \sigma/en$.

account for the residual carrier density, which leads to a finite minimum conductivity (or finite maximum resistivity).

In the presence of a magnetic field, the conductivity and resistivity acquire a tensorial form and can be translated into each other by $\hat{\rho} = \hat{\sigma}^{-1}$ (see Appendix B.2 and Figure B.2 therein). The following discussion is restricted to the perpendicular magnetic field case ($B = B_z$) and long channel geometries ($L/W > 1, \approx 3$). For the single carrier type regime, charges drift with the velocity v_D along the electric field lines in x-direction when no magnetic field is applied. When the magnetic field is turned on, they get deflected by the Lorentz force and accumulate on one side of the channel. This charge separation results in an electric field in y-direction whose force finally compensates the Lorentz force and the charge carriers drift again along the x-direction in equilibrium. The corresponding Hall voltage $U_H = E_y W$ can be measured between the (transverse) Hall contacts. With the condition $F_{el} = -F_L$ and $v_x = j_x/qn$ one can easily deduce the electric field component in the transverse direction

$$E_y = E_H = \frac{[r_H]}{qn} j_x B \,. \tag{2.39}$$

Using $I = j_x W$, the sign of the Hall voltage $U_H = [G] \frac{[r_H]}{qn} I B$ is determined by the charge carrier type for fixed current and magnetic field direction. In principle, the measured Hall voltage can also depend on a geometrical prefactor G which is, however, close to unity for our standard geometries ($G \approx 1$ for $L/W \approx 3$ [150]) and thus omitted in the following. The transverse resistivity is then given by measurable quantities $\rho_{yx} = R_{yx} = U_H/I$. The Hall coefficient (or Hall constant) is defined as

$$R_H \equiv \frac{E_y}{j_x B} = \frac{\rho_{yx}}{B} = \frac{[r_H]}{qn} \,. \tag{2.40}$$

We have so far assumed that all charge carriers move with the same velocity. At finite temperature, the carrier energies follow the Fermi Dirac distribution and the dominant scattering mechanism can be energy dependent. This requires a full treatment in Boltzmann kinetic theory which can be found in Refs. 150–152. In the single carrier type regime and for weak magnetic fields, this approach results in an additional factor, the so-called Hall scattering factor r_H, in the formula for the Hall coefficient. The latter has been already considered in all equations above. For

materials possessing spherical constant energy surfaces[26], the Hall scattering factor for low magnetic fields[27] is given by the ratio [150, 152]

$$r_H = \frac{\langle \tau^2 \rangle}{\langle \tau \rangle^2} \quad \text{with} \quad \langle \tau^m \rangle = \frac{\int E\, \tau^m(E)\, [-\partial f/\partial E]\, dE}{\int E\, [-\partial f/\partial E]\, dE}\ (m = 1; 2)\,, \tag{2.41}$$

where f is the Fermi-Dirac distribution and τ is the Boltzmann scattering time. The Hall scattering factor depends on the dominant scattering mechanism in the material and, thus, can be temperature and carrier density dependent, as shown for various materials [150, 154, 155]. For all known scattering mechanisms one finds $r_H < 2$ and typically $r_H \approx 1$ [150, 155]. The exact calculation of its value is a complex task and requires a profound model of the scattering mechanism [154–157]. For graphene, such theoretical calculations have not yet been performed and also no experimental evidence for significant deviations from $r_H \approx 1$ have been reported so far [1, 82, 142]. In Chapter 4.2 we will consider the Hall scattering factor when discussing asymmetries of the Hall coefficient values between the electron and hole side.

In experiment, it is common to measure the Hall coefficient R_H in order to obtain the carrier density. Together with the measured longitudinal resistivity ρ_{xx} (in the absence of a magnetic field), the so-called Hall mobility can be calculated from directly measurable quantities

$$\mu_H \equiv \frac{R_H}{\rho_{xx}} = R_H \sigma_{xx} = \frac{\rho_{yx}}{B\, \rho_{xx}}\,. \tag{2.42}$$

It is obvious that its value differs from the field effect mobility (Eq. 2.38) by the Hall scattering factor. Accordingly, the latter can be also defined as [150]

$$r_H = \frac{\mu_H}{\mu_{FE}}\,. \tag{2.43}$$

Note, this argumentation only holds when the gate factor β is calculated by knowing the exact dielectric thickness and the dielectric constant (or knowing the density from other physical effects, such as SdHO at low temperature). When the gate factor is determined from Hall measurements, the Hall scattering factor will be included implicitly in the derived gate factor. The Hall mobility also behaves differently when the CNP is approached. Since the measured R_H will become zero due to the coexistence of electrons and holes, the Hall mobility vanishes while the field effect mobility diverges. Because both definitions only hold at high density in the single

26 Otherwise, anisotropy and mixing effects have to be considered [150, 153].

27 Otherwise $r_H \to 1$ for $\mu^2 B^2 \gg 1$, that means for high magnetic fields and/or high mobility [150, 152].

carrier type regime, where we expect their values to agree very well, the onset of their deviation can serve as another estimate for the residual carrier density at the CNP.

The expression for the longitudinal resistivity ρ_{xx} remains independent of the magnetic field in the single carrier type regime (see Appendix B.2). Accordingly, no *classical*[28] magneto-resistance arises (for long channels) [110, 158].

Boltzmann transport theory for short and long range disorder potentials

The Boltzmann transport theory is a semiclassical approach to describe diffusive transport in a homogeneous system (in single carrier type case), which we restrict here further to two dimensions. The scattering events are described quantum-mechanically by transition matrix elements. They are calculated using Fermi's golden rule (in Born approximation) and the Pauli principle is considered with the Fermi-Dirac distribution. However, the theory neglects many-body interactions and quantum interference effects. The Boltzmann equation connects the disturbance in the equilibrium carrier distribution function caused by external fields with the equilibrium restoring effect of charge carrier scattering. A general introduction to the theory is provided in Refs. 150–152, 159 and its adaption to graphene can be found for example in Refs. 3, 28, 29, 32, 102. For the latter, the system is assumed to be in thermal equilibrium and homogeneous, such that the corresponding terms in the full Boltzmann kinetic equation are zero. Moreover, one considers the external fields to be weak such that the linear form of the Boltzmann equation can be used (first order approximation). Then the conductivity for graphene within the relaxation time approximation is given by

$$\sigma(\mathrm{E},\mathrm{T}) = \frac{\mathrm{e}^2}{2}\int \mathrm{dE}\ \mathrm{DOS}(\mathrm{E})\ \mathrm{v}_\mathrm{F}^2\ \tau(\mathrm{E},\mathrm{T})\,(-\partial \mathrm{f}(\mathrm{E})/\partial \mathrm{E})\,, \tag{2.44}$$

with the ideal DOS of graphene (Eq. 2.16) and the Fermi-Dirac distribution f(E) (Eq. 2.21). An alternative expression is $\sigma = (\mathrm{e}^2/\mathrm{h})(2\mathrm{E}_\mathrm{F}\,\langle\tau\rangle/\hbar)$ where $\langle\tau\rangle$ is the energy averaged finite temperature relaxation time, defined in Eq. 2.41 [32].

28 In short samples (Corbino geometry), the macroscopic current deflection leads to a geometrical magneto-resistance. Moreover, the so-called physical magneto-resistance also occurs in long samples. Like the Hall scattering factor, this can be explained semi-classically by the energy distribution of the charge carriers: The Hall field compensates the Lorentz force only for an average velocity, which is assumed equal in the Drude picture for all charge carriers. Most of the charge carriers, however, have slightly different velocities and are still deflected from their straight free transits [150]. This microscopic current deflection leads to the physical magneto-resistance which typically is very weak and negligible.

The transport relaxation time τ or rate τ^{-1} is the central ingredient in the Boltzmann transport theory. In contrast to the Drude model which relies on an ad-hoc scattering time and assumes the carrier to have a mass, the Boltzmann transport theory is also applicable to massless carriers and results in an explicit expression for the transport relaxation time. Its general form for the leading order Born approximation reads [28, 29, 160]

$$\frac{1}{\tau(\mathrm{E}_{\mathbf{sk}})} = \frac{2\pi}{\hbar}\sum_{\alpha}\int\frac{\mathrm{d}^2\mathrm{k}'}{(2\pi)^2}\,|\left\langle \mathrm{V}^{\alpha}_{\mathbf{sk},\mathbf{sk}'}\right\rangle|^2\,(1-\cos(\varphi_{\mathbf{kk}'}))\,\delta(\mathrm{E}_{\mathbf{sk}}-\mathrm{E}_{\mathbf{sk}'})\,, \quad (2.45)$$

where $\left\langle \mathrm{V}^{\alpha}_{\mathbf{sk},\mathbf{sk}'}\right\rangle$ is the matrix element of the α-kind scattering potential describing an arbitrary $\mathbf{k}\to\mathbf{k}'$ scattering process (inter- or intravalley) with angle φ. Inelastic interband processes ($\mathrm{s}\neq\mathrm{s}'$) are neglected [28]. The sum over α corresponds to the general form of Matthiessen's rule. To a good approximation, we can assume the total scattering rate to be given as the sum of the rates of each microscopic scattering process

$$\frac{1}{\tau_{\mathrm{tot}}} = \frac{1}{\tau_{\mathrm{C}}} + \frac{1}{\tau_{\mathrm{RSF}}} + \frac{1}{\tau_{\mathrm{phonons}}} + \frac{1}{\tau_{\mathrm{delta}}} + \ldots \quad (2.46)$$

One now has to identify microscopic scattering mechanisms, calculate their matrix element, and vary their parameters to fit experimental results. An overview and qualitative description of proposed mechanisms are given in the next subsection. Two exemplary mechanisms are also discussed in more detail at the end of this section.

From the definition of the Boltzmann conductivity (Eq. 2.44) follows that even when the scattering time for a certain mechanism is not explicitly temperature but energy dependent, a temperature dependence of the conductivity arises due to energy averaging $(-\partial \mathrm{f}(\mathrm{E})/\partial\mathrm{E})$. To be able to exclude the latter effect as well as phonon scattering mechanisms, one compares possible microscopic disorder mechanisms with experiments at low temperature ($\mathrm{T}\to 0$). For low temperatures and/or high densities ($\mathrm{k_B T}\ll \mathrm{E_F}$), Equation 2.44 simplifies to

$$\sigma(\mathrm{E_F},\mathrm{T}\to 0) = \frac{\mathrm{e}^2}{\mathrm{h}}\frac{2\mathrm{E_F}}{\hbar}\tau(\mathrm{E_F}) = \frac{\mathrm{e}^2\mathrm{v_F^2}}{2}\,\mathrm{DOS}(\mathrm{E_F})\,\tau(\mathrm{E_F}) = \mathrm{e}^2\,\mathrm{DOS}(\mathrm{E_F})\,\mathrm{D}. \quad (2.47)$$

The last part is the Einstein relation with the two-dimensional diffusion coefficient $\mathrm{D}=\mathrm{v_F}\,\mathrm{L_{mfp}}/2$ (with $\mathrm{L_{mfp}}=\mathrm{v_F}\tau$). For the above assumptions, Equation 2.20 holds and we can use $\mathrm{E_F}=\hbar\mathrm{v_F}\sqrt{\pi\mathrm{n}}$. The transport scattering time and the mean free path can then be directly calculated from the measured conductivity at the gate controlled (average) density $\mathrm{n}=\mathrm{n_g}$ according to

$$\tau(\mathrm{n}) = \frac{\mathrm{h}}{2\mathrm{e}^2\mathrm{v_F}\sqrt{\pi\mathrm{n}}}\,\sigma(\mathrm{n}) \quad \text{and} \quad \mathrm{L_{mfp}}(\mathrm{n}) = \mathrm{v_F}\,\tau(\mathrm{n})\,. \quad (2.48)$$

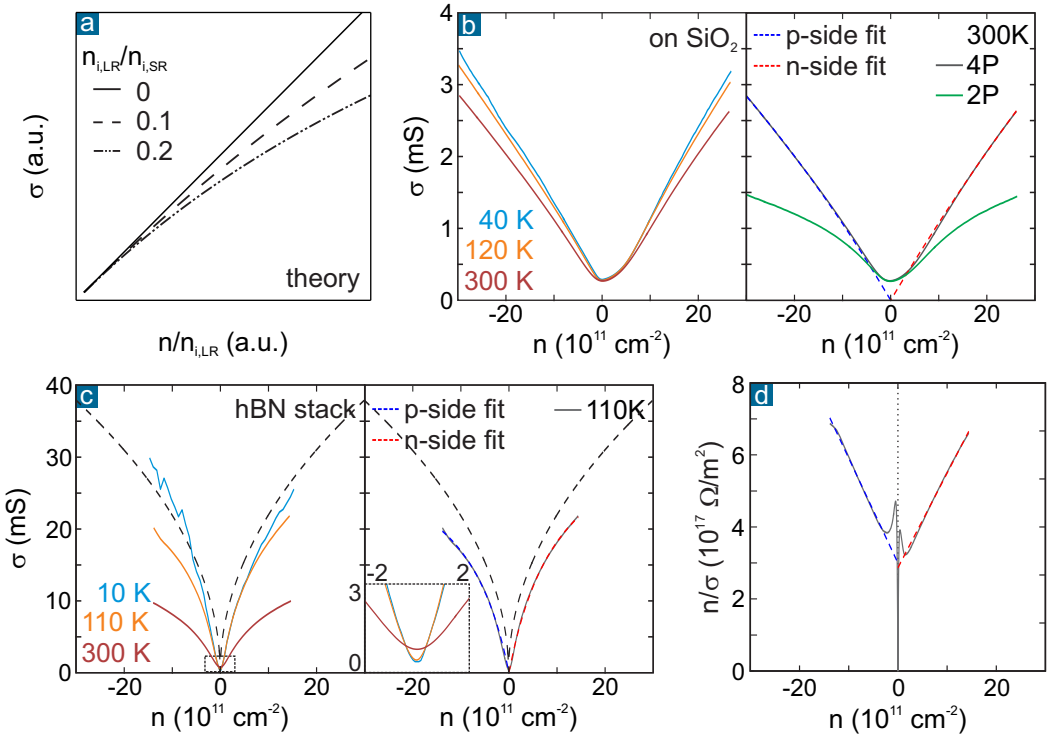

Figure 2.6. Long- and short-range contributions to the conductivity. a: Qualitative density dependence of the Boltzmann conductivity (for $T \to 0$) assuming long range screened charged impurity scatterers with density $n_{i,LR}$ and short range point defect scatterers with density $n_{i,SR}$ (modified from Ref. 32). b: Measured conductivity versus gate controlled (average) density for various temperatures for graphene on SiO_2 (left panel). The high density conductivity decreases with increasing temperature indicating phonon scattering. Right panel: Fit to the measured conductivity using Equation 2.51 ($\mu_{LR} = 0.7\,m^2/Vs$ and $\rho_{SR} = 50\,\Omega$). A comparison between two (green line) and four terminal measurements indicates the importance of four terminal measurements to avoid the contact resistance contribution. c: Same analysis as in panel b, but for hBN-encapsulated graphene (note the different y axis scaling), fit results: $\mu_{LR} = 21\,m^2/Vs$ and $\rho_{SR} = 27\,\Omega$. Black dashed line corresponds to the ballistic conductivity limit at high density for the given sample width (see Eq. 2.71 in Section 2.4). d: Same data and fit as in c, but plotted differently, as given by the second part of Eq. 2.51.

And the mobility $\mu = \sigma/e\,n$ (Eq. 2.37 and 2.38) is given by

$$\mu(n) = \frac{\sigma(n)}{e\,n} = \frac{2e}{h}\sqrt{\frac{\pi}{n}}\,v_F\,\tau(n) = \frac{2e}{h}\sqrt{\frac{\pi}{n}}\,L_{mfp}(n)\,. \tag{2.49}$$

We now consider scalar scattering potentials in the low temperature limit which do not couple to the pseudo spin part of the wave function (see Section 2.1.5 and Refs. 7, 29, 160). The Fourier transform of such potentials is given by a function $V(q)$ with $q = |\mathbf{k} - \mathbf{k}'|$. As an example for a long range potential, we take into account (screened) charged impurities, also referred to as Coulomb scatterers, which are here, for the sake of simplicity, assumed to be located exactly at graphene's surface. The Fourier transform then reads $V^C(q) = 2\pi e^2/\kappa q$. Moreover, we consider (weak) short range defect scatterers which are described by zero range (delta) potentials and $V^S(q) = V_0 (= \text{const.})$. As shown in Section 2.1.5, the pseudospin part for intravalley scattering results in the chiral factor $w(\varphi) = \cos^2(\varphi/2) = (1 + \cos\varphi)/2$ and the scattering rate can be written as

$$\frac{1}{\tau(E)} = \frac{\pi}{\hbar} \sum_{\alpha,k'} n_{imp}^{\alpha} \left| \frac{V^{\alpha}(q)}{\epsilon(q)} \right|^2 (1 - \cos^2\varphi)\, \delta(E - E_{k'})\,, \tag{2.50}$$

where $\epsilon(q)$ is the RPA dielectric screening function (Eq. 2.25) and n_{imp} is the impurity density of each scatterer type. The angular dependence is depicted in Figure 2.2b indicating the short and large angle scattering suppression for such potentials [102]. Due to the unusual screening behavior in graphene (see Section 2.2.3), the functional dependence of bare and screened Coulomb interaction on the carrier density remains the same apart from a multiplicative constant. One can then show that the scattering times (or equivalently the corresponding mean free paths) are proportional to $\tau_C \propto k_F \propto \sqrt{n}$ and $\tau_S \propto 1/k_F \propto 1/\sqrt{n}$, respectively [3, 7, 32]. Hence, the Boltzmann mobility (Eq. 2.49) is constant for charged impurities whereas it scales as $1/n$ for the zero range defect scatterers which are considered here as example for short range potentials. Equivalently, Equation 2.47 reveals that σ_C depends linearly on the density and σ_S is density-independent [32].

Using the general form of Matthiessen's rule (Eq. 2.46), the total conductivity is composed of a linear term corresponding to long range disorder potentials (as outlined above for the case of Coulomb scatterers) and a constant term corresponding to short range potentials (as outlined above for the more specific case of zero range defect scatterers). Using a more general notation (subscripts LR and SR) the conductivity is given by

$$\frac{1}{\sigma} = \rho = \rho_{LR} + \rho_{SR} = \frac{1}{e\,\mu_{LR}\,n} + \rho_{SR} \;\leftrightarrow\; \frac{n}{\sigma} = \frac{1}{e\,\mu_{LR}} + \rho_{SR}\,n\,. \tag{2.51}$$

This equation was used in many publications to capture the functional dependence in experimental results and determine the non-linearity of the conductivity [30, 39, 161, 162]. We note that in the literature often the long range mobility μ_{LR} is stated as "the" mobility in graphene. However, this is only in agreement with

the field effect mobility definition (Eq. 2.38) when long range disorder scattering prevails over a broad density range, i.e. for lower quality devices. The separation into long range and short range type disorder is also motivated by the fact that only[29] long range disorder can explain electron hole puddle formation near the CNP for temperatures below the disorder equivalent temperature scale (see Section 2.2.3). Finally, one can estimate the corresponding mean free path $L_{LR} = \hbar\sqrt{\pi n}\mu_{LR}/e \propto \sqrt{n}$ and $L_{SR} = \hbar\sqrt{\pi n^{-1}}/(e^2\,\rho_{SR})$ from the fitting parameters (μ_{LR} and ρ_{SR}) used in Eq. 2.51 [3, 162].

At finite temperature, all energy-dependent scattering mechanisms which are not explicitly temperature dependent will become (weakly) temperature dependent due to the energy averaging in Eq. 2.44. Furthermore, for scattering potentials that are screened by the dielectric function, the temperature dependence of this function has to be included (that means $\epsilon(q, T)$). Theoretical studies have done this for charged impurities in Refs. 29, 102, and additionally for (acoustic) phonons and short range δ-disorder in Refs. 80, 149, 160. By doing so, temperature dependent charged impurity scattering is used in an effective medium approach to develop a universal disorder model (see Section 2.3.3 and Refs. 80, 109).

2.3.2. Charge carrier scattering and disorder mechanisms

Several microscopic scattering mechanisms have been proposed to contribute to momentum relaxation in graphene (see Table 2.2). An overview of these mechanism is for example given in Refs. 3, 7, 9, 115, 163. In this work we restrict ourselves to pristine graphene with very low *intrinsic* disorder level, that means a very low density of vacancies and lattice defects. This is realized in exfoliated and single-crystalline CVD grown graphene. By doing so, we exclude scattering at grain boundaries in polycrystalline graphene, defects related to the formation of covalent bonds to the substrate as in epitaxial graphene, and intentionally induced vacancies (described in Refs. 9, 164).

Moreover, in the following, we assume *extrinsic* impurities to not be strong enough to generate extra structure in the DOS [7]. That means, for instance, they do not

29 Based on the scattering behavior alone, there is also short range disorder that can lead to a linear density dependence of the conductivity and correlated long range disorder that might explain the sublinearity at high density (see next section). However, the short range disorder fails to explain the electron hole puddle landscape and the correlation of long range disorder scatterers cannot be explained by a common microscopic picture on all substrates. Hence, we stick here to the well established separation in long range (linear dependence) and short range (constant conductivity) disorder.

bond covalently to the graphene lattice. Otherwise, such covalent bonds would inflict a hybridization modification from sp^2 to sp^3. If this were the case in real devices, these defects would be described by strong short range disorder potentials which modify substantially the band structure of graphene and add additional structure to the DOS. Several theoretical studies have discussed such strong defects, also referred to as resonant impurities or midgap states in the literature [7, 165–167]. It was found that the linear density dependence of the conductivity and its crossover to a sublinear behavior at higher densities can be explained by resonant scatterers. However, this disorder model exhibits several deficiencies: Its microscopic origin is unclear for the considered pristine graphene case (even though hydrocarbons were suggested as possible origin of resonant scattering [167, 168]). A substantial modification from sp^2 to sp^3 would be indicated by the presence of a notable D peak in graphene's Raman spectrum. However, this peak is virtually absent for pristine graphene samples, except for the edges (see Chapter 3). Furthermore, the model of strong short range scatterers fails to explain the saturation of the conductivity to a minimum at the CNP in experiment and no theory is available to describe quantitatively the electron hole puddle landscape [7, 169]. Finally, strong short range scattering would imply that intervalley scattering is dominant. This is contrary to all weak localization experiments conducted so far on different substrates which point to intravalley scattering instead (see Appendix B.1). Hence, we will not consider resonant scatterers[30] in the following discussion but emphasize their importance for

30 After this chapter has been finished, Halbertal and coworkers (Ref. 170) found, using scanning nanothermometry performed at 4 K, that resonant scatterers (localized states caused by atomic-scale defects) govern the energy dissipation of current generated hot electrons in very clean hBN-encapsulated graphene devices and possibly limit the mobility in pure graphene. The authors argue that such defects enhance (locally) the inelastic electron-phonon scattering resulting in efficient energy dissipation pathways (via ballistic phonons). Vacancies or adatoms that are covalently bonded to a single carbon atom, e.g. individual hydrogen adatoms, constitute possible candidates for such resonant defects (which implies the formation of a local sp^3 orbital). These atomic defects are found to be very rare in the bulk, as one expects for exfoliated flakes, but abundant along the plasma etched edges of the graphene-hBN heterostructure. Hence, the main part of the energy dissipation is found to take place at the edges of the high quality device at low temperature (when the momentum relaxation mean free path surpasses the device dimensions over a wide density and temperature range). Accordingly, the observed resonant dissipation within the mesoscopic device, including its edges, opposes the common Landauer picture where the energy dissipation is expected to occur at the electrodes (far away from the measured sample). Moreover, Halbertal and coworkers report that the dissipation at the edges remains dominant over the whole studied density range, including the CNP. Considering the device's high quality and the temperature of 4 K, charge transport is obviously ballistic at higher densities, in agreement with earlier studies which also find clear evidence of boundary scattering (see Section 2.4). However, all (electrical transport) experiments clearly observe a diffusive regime at the CNP (minimum conductivity) and the formation of a electron hole puddle landscape (even at low temperature and for ultra high quality devices). Long range type disorder, such as charged impurities, explains these findings while a theoretical framework

mechanism	$\sigma(n)$	type	T dep.	microscopic origin
random charged impurities [29]	$\propto n$	LR	$\langle\tau\rangle$, $\epsilon(q,T)$	impurities close to or at surface
correlated charged imp. [174]	sublinear	LR	$\langle\tau\rangle$, $\epsilon(q,T)$	correlated impurities after annealing
random strain fluct. [18, 131]	$\propto n$	LR	not yet addressed	in-plane and out-of-plane deformations
ripples [133]	$\propto n^{2H-1}$	SR/LR	not yet addressed	corrugations
resonant impurities [165]	$\propto n \ln^2(\sqrt{n})$	SR	not yet addressed	unclear in pristine devices
zero range δ [32, 160]	const.	SR	$\langle\tau\rangle$, $\epsilon(q,T)$	unclear
acoustic phonons [78, 175]	$\rho(T)$	SR	$\rho \propto T$	longitudinal mode
flexural phonons [130]	$\rho(T)$	SR	$\rho \propto T^2$	only suspended
optical phonons [78]	$\rho(T,n)$	LR	activated	intervalley (A_1') mode
remote interfacial phonons [90, 176]	$\rho(T,n)$	LR	activated	in polar substrates

Table 2.2. Momentum relaxation mechanisms in monolayer graphene. When the mean free path is larger than the sample dimensions, boundary scattering may become the main mechanism for momentum relaxation. The conductivity in this ballistic regime is then proportional to the channel width, $\sigma_{\text{ball}} \propto W\sqrt{n}$ (see Eq. 2.71). For devices with extremely low disorder, electron-hole collisions can prevail for intermediate temperatures at the CNP. They then limit the charge carrier transport (see Section 2.4).

experiments with intentionally induced strong defects or adsorbates [168, 171, 172], and their contribution to the scattering at high densities for graphene on non-conventional substrates such as $SrTiO_3$ [169, 173].

based on resonant scatterers is lacking. Therefore, the observation of Halbertal et al., that not Joule heating but energy dissipation at the edges is dominating also at the CNP, is quite surprising and stresses the importance of taking into account the finite dimensions of today's heterostructure devices.

The main long range disorder mechanisms leading to $\sigma \propto n$

In the low temperature limit[31], the long range disorder potential, causing the spatial density fluctuations near the CNP, as well as the main scattering mechanism at low density have the same microscopic origin [7, 18]. In the following discussion we address in more detail the dominant long range disorder mechanisms. In Section 2.2.5, charged impurities and random strain fluctuations were already identified as those mechanisms that agree best with the experimental findings obtained from local probe experiments. Here, we complete this picture by focusing on the results obtained from electrical transport measurements together with the theoretical description of both mechanisms. The transport measurements are carried out at low temperature such that phonon scattering can be neglected.

Randomly distributed, RPA-screened charged impurities are the most popular source for both charge carrier scattering and puddle formation [7, 29, 31]. The theoretical framework is well developed and not only covers the linear density dependence of the conductivity in the single carrier type regime, but also provides a quantitative description of the conductivity (minimum) in the two carrier type regime around the CNP (when the Boltzmann transport theory for charged impurity scattering is implemented in the effective medium theory, as discussed in Section 2.3.3). Charged impurities are assumed to be located either close to the substrate surface or directly at the graphene lattice (e.g. trapped adsorbates and dangling bonds). The Boltzmann conductivity for such screened long range Coulomb scatterers involves three modeling parameters: the impurity density n_{imp}, the average distance of the impurities to the graphene plane d (often set to zero, and only n_{imp} is adjusted in order to fit the experimental results) and the relative strength of the Coulomb interaction defined by the parameter r_s (see Eq. 2.10). In the low temperature limit and in the single carrier type regime (for a homogeneous system with density n), the Boltzmann conductivity limited by charged impurity scatterers (assuming $d = 0$) is given by [7, 29, 32, 99]

$$\sigma_c(n) = A[r_s]\frac{e^2}{h}\frac{n}{n_{imp}} \quad \text{with} \quad \mu_c = A[r_s]\frac{e}{h}\frac{1}{n_{imp}}\,. \tag{2.52}$$

Here, $A[r_s]$ is an auxiliary function that can be written in an analytical form [29, 99]. It is a result of the RPA theory which is employed to model the screening of Coulomb

31 Strictly speaking, low temperature here refers to a temperature below T_c (as defined in Eq. 2.29). Above that temperature, thermal activation of carriers substantially smoothens the electron hole puddle landscape and optical phonon scattering can prevail at low density. Below that temperature, the temperature dependence of the conductivity near the CNP follows a universal disorder limited curve (as shown for charged impurity disorder in Ref. 80 and discussed in more detail in Section 2.3.3).

interactions in graphene (as already discussed in Section 2.2.3 where the dielectric screening function of graphene in the low temperature limit $\epsilon_{RPA}(q)$ was given in Eq. 2.25). Since r_s is density-independent for graphene's linear energy dispersion, the conductivity depends linearly on the density and, equivalently, the mobility μ_c is a constant with respect to density. However, r_s depends on the dielectric environment of graphene which is characterized by the effective dielectric constant ($r_s \propto 1/\kappa$). $A[r_s]$ is found to decrease monotonically as r_s becomes larger [29, 99]. Hence, an increase of the dielectric constant of the environment is expected to cause an increase of the mobility μ_c. Moreover, a reduction of the impurity density n_{imp} (including both increased distance from the surface and the strength) will also enhance the mobility.

The theoretically predicted dependence on the impurity density and the dielectric environment has been verified in several low temperature transport experiments[32] in which these parameters were tuned systematically [30, 82, 91, 149, 177–179]. Accordingly, the substantially lower intrinsic impurity density of hexagonal boron nitride, measured independently by local probe techniques (see Section 2.2.5), explains why the mobility of graphene on hBN is much larger than on SiO_2, although the dielectric constants are similar [160, 180]. Moreover, the self-cleansing effect leads to atomically clean interfaces free of contaminants when graphene is encapsulated (improving further the mobility). Finally, the ultra high mobility at low temperature in case of suspended graphene is also attributed to an extremely low charged impurity density as a result of current annealing in vacuum (while the dielectric constant is the lowest in the absence of a substrate, not favoring a high mobility). To conclude, the most critical requirement for high mobility is a low impurity density. An optimization of the dielectric background is far less efficient.

At low temperature and low density, in the two carrier type regime, the system is spatially inhomogeneous (caused by the same disorder mechanism as the one responsible for scattering). The Boltzmann transport theory, presuming a homogeneous system, does not apply anymore. Experiments find the conductivity to

32 Other studies have found that the low temperature mobility is not enhanced by placing graphene on high-κ substrates [135, 173]. They then argued that charged impurity disorder may not be the dominant scattering mechanism. However, it is difficult to control selectively the dielectric environment by changing the substrate for graphene without also unintentionally changing the impurity density (as argued in Ref. 169). The impurity density can easily vary both due to the different intrinsic charged impurity density of those substrates (not double-checked by these experiments) as well as due to a different adsorbate density (stemming from modifications in the device processing routine, e.g. a polymer-based transfer of the graphene flakes, or related to different functional groups of the substrate surface). A later published study was able to verify the predicted κ-dependence by varying the dielectric environment of suspended graphene without involving a change in the substrate material [177].

approach a minimum $\sigma_{\min(T\to 0)}$. The self-consistent approximation theory (SCA), first introduced for charged impurity disorder in Ref. 99 and generally discussed in Section 2.2.3, assumes the system to be homogeneous with a disorder related density n_*^{SCA}. The theory then determines this value self-consistently such that the Boltzmann conductivity for this density $\sigma_B(n_*^{SCA})$ equals the experimentally measured minimum conductivity at the CNP. The measured conductivity in the low density regime ($n < n_*^{SCA}$) is then approximated by a constant plateau $\sigma_B(n_*^{SCA})$, while for $n > n_*^{SCA}$ the system is considered homogeneous and the Boltzmann conductivity in the single carrier regime $\sigma_B(n) \propto n$ applies (see also panel a of Figure 2.7 and Ref. 7). The general equation for the self-consistent residual density (Eq. 2.26) transforms for charged impurity disorder into [29, 99]

$$n_{*,c}^{SCA} = 2\, n_{imp}\, r_s^2\, C^{RPA}[r_s, a = 4d\sqrt{\pi n_{*,c}^{SCA}}]\,. \tag{2.53}$$

Here, the auxiliary function C^{RPA}, a result of the RPA theory, accounts for self-screening effects (calculated in Ref. 99). As we already discussed in Section 2.2.3, the SCA density can be estimated from low temperature transport[33] measurements (corresponding to $n_{t,0}(T \to 0)$ in our notation), whereas the mobility can be determined independently from the linear increase of the conductivity away from the CNP. Comparing many different devices (on hBN and SiO_2) with varying levels of disorder, one finds that the residual density at the CNP and the mobility away from the CNP are correlated at low temperature (see Ref. 18 and Figure 3.8). Moreover, they seem to follow a universal curve, independent of the substrate. It is argued in Ref. 18 that this finding supports the picture of dominant random strain fluctuations disorder. However, to our knowledge, it has not yet been checked by theorists whether this correlation could also be explained in the charged impurity model[34]. Another advantage of the charged impurity picture is that it explains self-consistently the initial doping level (the shift of V_{CNP}) observed in several, typically non-encapsulated devices [99]. This is not captured by the random strain fluctuations picture.

33 Alternatively, local probe experiments allow a direct determination of the root-mean-square density fluctuations n_{rms}. Assuming a Gaussian distribution one obtains $n_*^{SCA} = n_{rms}/\sqrt{3}$ [109].

34 Combing Eq. 2.52 and 2.53 one obtains $1/\mu_c = (2\, A[r_s]\, r_s^2\, C[r_s, a(n_{*,c}^{SCA})])^{-1}\, (h/e)\, n_{*,c}^{SCA}$. For $\kappa \approx 4$, one gets $r_s \approx 0.8$ and $A[r_s] \approx 20$ [99]. This gives us $1/\mu_c = 0.039 \cdot C[r_s, a(n_{*,c}^{SCA})]^{-1} (h/e)\, n_{*,c}^{SCA}$ (compare to Eq. 2.54). Since $C[r_s, a(n_{*,c}^{SCA})]$ depends itself on the residual density (via the parameter a), the functional dependence would not be exactly linear to $n_{*,c}^{SCA}$. Moreover, due to the overall dependence on r_s, the relation is likely not universal on all substrates with different dielectric constants. However, the dielectric constants of hBN and SiO_2 (for the devices studied so far) are quite similar and even if r_s was different for another substrate type, an exact calculation of the change in $(A[r_s]\, r_s^2\, C[r_s, a])$ has to be performed (which has not yet been done, to our knowledge).

The temperature dependence of the charged impurity disorder limited conductivity arises from the temperature dependence of the dielectric screening function $\epsilon(q, T)$ as well as from the finite temperature energy averaging in the Boltzmann theory (see Eq. 2.44). The full calculations are provided in Ref. 102. At higher densities, in the single carrier type regime when the Boltzmann theory can be applied, the charged impurity limited conductivity slightly decreases, in good approximation linearly, as the temperature increases (corresponding to weak metallic behavior). At low density, the temperature dependence of the Boltzmann conductivity is complex and can show both insulating and metallic behavior. However, as we will discuss in Section 2.3.3, an effective medium theory (EMT) better captures the inhomogeneous system near the CNP (when the Boltzmann theory cannot be employed). In this case, the transport inside the puddles is described by the Boltzmann theory including the above discussed temperature dependence for the charged impurity limited Boltzmann conductivity.

In summary, the vast majority of transport experiments, in particular those which tune the main parameters (n_{imp}, r_s) in a controlled fashion, can be successfully described over the whole density range by the charged impurity disorder model (in the low temperature limit in the absence of phonon scattering). Moreover, the model is appealing since it depends only on two modeling parameters (n_{imp} and r_s) and its main results are typically given in an analytical form. The same conclusion we already made for the local probe experiments where the charged impurity disorder theory nicely captures the puddle formation (see Section 2.2.5). We also argued that the charged impurity picture may explain the initial doping found in dirtier samples as well as the ultimate improvement of the device quality when using thin graphitic gates (i.e. multilayer graphene), substituting the conventional silicon back gate, in hBN-encapsulated devices (which points to the importance of remote charged impurities). However, an unsolved issue of the charged impurity model arises from weak localization experiments (see Appendix B.1). They reveal that intravalley symmetry breaking scattering is dominant in the low temperature limit and for a density range where the conductivity depends linearly on the density (that means where the charged impurity model should clearly be applicable). As discussed in Section 2.1.5, charged impurity disorder results in symmetry conserving intravalley scattering which disagrees with these experimental findings.

Random strain fluctuations (RSF) have been proposed recently as another dominant disorder mechanism at low temperature and for lower densities [2, 18, 131, 181]. As shown in the following, both the linear density dependence of the conductivity as well as the residual density at the CNP, $n^{SCA}_{*,RSF} = n_{t,0}(T \to 0)$ (the plateau width of the conductivity minimum), can be explained by this mechanism. Strain may originate from in-plane and out-of-plane deformations. The latter are induced by

the roughness of the underlying substrate causing graphene to conform to the corrugations [127, 128, 182]. Early theories considered only these out-of-plane deformations, also referred to as ripples[35]. A more advanced theoretical model, provided by Couto and coworkers in Refs. 18, 131, accounts additionally for in-plane strain. Such in-plane strain can arise even on atomically flat substrates when graphene suffers almost no out-of-plane deformations but its carbon atoms may experience a *random* distribution of in-plane forces, due to the interaction with imperfections in the substrate or with adsorbates trapped between graphene and the substrate[36].

The random strain fluctuations (both out-of-plane and in-plane) are described by introducing a scalar potential *and* a vector potential (a gauge potential) in the Hamiltonian [18, 131]. Both have long range character and contribute to charge carrier scattering. The scalar potential is subject to screening which is accounted for by graphene's dielectric function $\epsilon^{\mathrm{RPA}}(q)$ (as for charged impurities). The vector potential acts like a random pseudomagnetic[37] field whose effect cannot be screened. The calculations performed by Couto and coworkers suggest that the scattering rate related to the gauge potential is at least one order of magnitude larger than the

35 The focus on these ripples (the microscopic corrugations of a graphene sheet) was also motivated by the observation of intrinsic ripples in freestanding graphene revealed in room temperature TEM experiments [16]. While the ripples in freestanding graphene are attributed to flexural phonons (limiting drastically the room temperature mobility in these suspended devices, as discussed in Refs. 17, 130 and Chapter 3), the charge carrier scattering on (static) microscopic corrugations in substrate supported graphene was mainly addressed in Refs. 3, 133. The authors find the density dependence of the conductivity to be determined by the type of the ripple correlations. These are parametrized as $\langle [h(r) - h(0)]^2 \rangle = r^{2H}$, where the value of H is related to the origin of the ripples [3]: $2H = 1$ implies short-range correlations (when graphene conforms to the substrate), while $2H = 2$ corresponds to long-range correlations (when graphene adheres only loosely to the substrate due to van der Waals forces, constituting a thermally excitable membrane). AFM measurements suggest $2H \approx 1$ but > 1 [3, 128, 182]. Then, $\sigma(n) \propto n^{2H-1}$ and the charge carrier scattering mimics short-range disorder, that means $\sigma(n) \approx$ constant. Accordingly, out-of-plane deformations treated in the approach of Ref. 133 alone cannot explain the linear density dependence but constitute a possible candidate for the constant conductivity limit at high density.

36 In this context, we emphasize the difference between these random (in-plane) strain fluctuations and the periodic ones occurring only for aligned crystallographic planes due to the incommensuration between graphene and hBN (which results in the emergence of a Moiré superlattice, as shown in Section 2.1.6)

37 A descriptive picture, why the effect of the gauge potential can be compared to a pseudomagnetic field, is given in Ref. 2: Let us consider a straight boundary between an unperturbed and a perturbed region with elastic strain along one bond direction of graphene's lattice. The modified hopping in the perturbed region leads to a displacement of the Dirac cones (K and K') in opposite directions (in k-space). In real space, this corresponds to a deflection of the charge carriers traveling across the boundary since both energy and the momentum parallel to the boundary are conserved. Accordingly, the trajectory of the charge carriers changes at the boundary as if there was a magnetic field applied perpendicular to the plane.

rate related to the scalar potential. As a consequence, the dominance of the gauge potential suggests that placing graphene on high-κ substrates should not enhance the mobility (since a better screening does not weaken this vector potential). This might[38] explain why the low temperature mobility for graphene on $SrTiO_3$ does not exceed that on SiO_2 [173]. In contrast to the scattering, the self-consistent residual density at the CNP $n^{SCA}_{*,RSF}$ is determined by the scalar potential *only* and can be derived from the general SCA equation (Eq. 2.26).

The full expressions for the density-independent mobility and the SCA residual carrier density in the random strain fluctuations model are bulky but can be given in an analytical form, as shown for the low temperature limit in Ref. 18. The main findings are that μ_{RSF} and $n^{SCA}_{*,RSF}$ are related and only depend on graphene's intrinsic properties, such as the electron-phonon coupling strength and the elastic properties characterized by the Lamé coefficients. Accordingly, the substrate only determines the magnitude of the random strain fluctuations in graphene's lattice but has no further impact (in this model). The higher mobility on hBN compared to SiO_2 is then explained by the much smaller surface roughness of the hBN flake (see Chapter 3). Apart from using substrates with a generally smoother surface, it is impossible to control precisely the magnitude of strain fluctuations in a certain device. However, it is possible to study a large ensemble of devices where the magnitude of the strain fluctuations statistically differs from sample to sample (see Figure 3.8). By doing so, Couto and coworkers found the correlation between the mobility (determined as the linear slope of the conductivity away from the CNP) and the residual carrier density to be described best by

$$\frac{1}{\mu_{RSF}} = 0.118 \, \frac{h}{e} \, n^{SCA}_{*,RSF} \, . \tag{2.54}$$

They claim that the value of this pre-factor is in agreement with their theory when assuming reasonable values for the Lamé coefficients and the electron-phonon coupling strength (as fitting parameters). The use of different substrate materials (mainly hBN and SiO_2) also supports the picture of a universal correlation.

Moreover, the random strain fluctuations theory is backed by the analysis of weak field magneto-resistance curves at low temperature which are governed by quantum interference effects (conducted in the same study of Couto and coworkers). As aforementioned, these experiments suggest a dominant intravalley symmetry-breaking mechanism which can be explained by the gauge potential acting as a pseudomagnetic field and coupling to the pseudospin [18, 131]. In other words, random strain

38 However, it can be also argued that the impurity density might increase and compensate the beneficial effect of better screening when assuming the charged impurity model, since these experiments do not explicitly account for that problem [169].

fluctuations break the sublattice symmetry in graphene. Finally, the picture of strain fluctuations is completed by Raman experiments [18, 123, 124]. The width of the 2D peak in the spectrum is related to nanometer-scale strain variations (averaged over the focal spot size) [18, 123]. It becomes more narrow for graphene on atomically flat hBN than on SiO_2 and hence can serve as a non-invasive tool for device pre-characterization (see also Chapter 3). Recently, several other groups, including us, provided additional experimental evidence that random strain fluctuations are the main source of disorder in high quality, hBN-encapsulated graphene devices (see Refs. 183–185 and Section 3.1.5).

In summary, the results of the reported transport experiments do not alter the conclusion we had made already in Section 2.2.3 when discussing local probe experiments. Both charged impurities and random strain fluctuations become more or less important depending on the substrate type, encapsulation, and the processing routine. Random strain fluctuations, together with remote charged impurities, set the ultimate limit on atomically flat substrates with low intrinsic charged impurity density, such as hBN. In contrast, charged impurities play the dominant role on the opposite type of substrate, such as SiO_2, and in the presence of adsorbates (which then also explains the initial doping). All in all, the theoretical framework of charged impurity disorder is (much) more mature than for the random strain fluctuations picture. For instance, the temperature dependence of the conductivity limited by random strain fluctuations has not yet been addressed. Moreover, a detailed picture of the puddle landscape in this framework, such as length correlations, has not yet been elaborated, on except for the DFT based calculations in Ref. 108 (see also Section 2.2.5).

Scattering mechanisms leading to constant σ and phonon scattering

At all temperatures, the dependence of the conductivity on the density becomes increasingly sublinear as the density increases (see Figure 2.6c). For devices with low extrinsic disorder, the conductivity even tends to approach a maximum at higher temperatures, e.g. at room temperature. Moreover, this maximum conductivity is substantially temperature dependent. This points to phonon scattering limiting the mobility at higher temperature (and higher density). However, the origin of the sublinearity at low temperature is unknown and still under debate [174].

In the previous section we have introduced the Boltzmann model for long range and short range scattering (in the low temperature limit) where an ad-hoc zero range scattering mechanism was postulated to describe the conductivity saturation (the

sublinearity) [32]. In real devices, the microscopic origin of such zero range scatterers is unclear, especially since intrinsic lattice defects are very rare. Other theories invoke correlations among charged impurities, e.g. as a result of annealing during device processing, or an additional contribution of resonant scatterers in order to explain the sublinearity at low temperature [169, 174, 186]. Ripples (out-of-plane deformations) induced by the substrate surface morphology were also found to mimic short range scattering leading to a constant conductivity [133, 182]. However, all these explanations are not applicable to high quality devices where these proposed microscopic mechanisms are mostly absent. Yet these devices show the most pronounced sublinearity. Furthermore, one has to take into account the typical device dimensions (which are on few micrometer scale). At low temperature, due to the large mean free path (high mobility) in such low disorder devices, a crossover to ballistic transport at higher densities is very plausible. Then, boundary scattering dominates leading to $\sigma(n) \propto \sqrt{n}$ (as discussed in Section 2.4 and shown in Figure 2.6c). This square root dependence can be easily confused with the sublinearity arising in the Boltzmann transport theory with long range or short range scatterers.

Phonon scattering in graphene is relatively weak compared to other semiconductors ($\omega_{LO}^{semic.} \ll \omega_{LO}^{graphene}$). In semiconductors, polar optical phonon scattering typically limits the room temperature mobility (far) below $1\,m^2/Vs$, e.g. in GaAs to $2000\,cm^2/Vs$ [175]. The main reasons for weak phonon scattering in graphene are a weak electron-phonon coupling strength and an extraordinary Debye temperature θ_D of about 2800K due to the strong sp^2 bonds and light carbon atoms making graphene an exceptionally stiff material with few thermally excited phonons at room temperature [187]. First theories therefore considered only low energetic *longitudinal acoustic phonons* as dominant carrier lattice interaction [175]. Since graphene's Fermi surface is very small, only small-q phonons can scatter electrons, and they are all populated above the Bloch-Grüneisen temperature $\theta_{BG} \ll \theta_D$. In Boltzmann theory the density-independent resistivity due to longitudinal acoustic phonon scattering above[39] θ_{BG} reads [90, 175]

$$\rho_{LA} = \frac{h}{e^2} \frac{\pi^2 D_A^2 k_B T}{2 h \rho_s v_s^2 v_F^2} \propto T\,, \tag{2.55}$$

where $\rho_s = 7.6 \cdot 10^{-7}\,kg/m^2$ is the 2D mass density, $v_s = 2.1 \cdot 10^4\,m/s$ the velocity of sound, $v_F = 1 \cdot 10^6\,m/s$ the Fermi velocity, and $D_A \approx 18 - 21\,eV$ the acoustic deformation potential. The value of this last parameter is typically adjusted in order

39 Below θ_{BG} the resistivity scales as $\rho_{LA} \propto T^4$ and the Bloch-Grüneisen temperature is found to become larger as the carrier density is increased [87, 187]. However, for the highest densities that can be reached by electrostatic gating θ_{BG} is always smaller than 10 K (that means not relevant for our work).

to conform the theory to experimental data [10, 90, 175]. Accordingly, if all other scattering mechanisms are eliminated, the residual resistivity at room temperature is about $40 - 50\,\Omega$, corresponding to a record mobility of $10\,\mathrm{m^2/Vs}$ at $\mathrm{n} = 10^{12}\,\mathrm{cm^{-2}}$ (note, since $\rho_{\mathrm{LA}} = \text{constant}$, $\mu_{\mathrm{LA}} \propto 1/\mathrm{n}$). This intrinsic mobility limit at room temperature, set by acoustic phonon scattering, can be reached in high quality, hBN-encapsulated devices, as we will discuss in Chapter 3.

When graphene is suspended, low energetic *flexural phonons* contribute significantly to carrier scattering ($\rho_{\mathrm{flex}} \propto \mathrm{T}^2$) [130]. This limits the room temperature mobility of suspended devices to only $1\,\mathrm{m^2/Vs}$ while the low temperature mobility is extremely high (due to the extremely low extrinsic disorder level). The rather poor room temperature mobility is comparable to the value obtained for devices on SiO_2 with high disorder (where the mobility is limited by charged impurity disorder both at room and low temperature). Although flexural phonon scattering can also be suppressed by straining the suspended devices, this additional scattering mechanism and the substantial technical difficulties in producing such fragile devices hamper the realization of suspended graphene in commercial applications.

Experiments on graphene residing on different substrates reveal that the resistivity increases much stronger above 200K than predicted by the LA-phonon model [10, 39, 90, 176, 188]. The deviation is larger at low densities and its magnitude cannot be alternatively explained by the weak temperature dependence of charged impurity scattering. *Remote phonon scattering* by polar modes which are thermally activated in polarizable substrates were proposed as a possible cause for this deviation [90, 189, 190]. The coupling between charge carriers in graphene and the long range field associated with surface phonons in the dielectric substrate are stronger at low carrier density due to weaker screening. The scattering rate at a certain temperature is determined by the activation energy of the optical phonons in the substrate and the coupling strength to this specific mode. Both parameters are substrate specific and surface phonon scattering in oxides has been identified as a crucial detrimental factor to mobility at room temperature, becoming even more important in high-κ dielectrics such as HfO_2 [176, 191]. Again, hBN turns out to be a nearly ideal substrate contributing a negligible remote phonon scattering induced resistivity of about 3Ω at room temperature [190].

Surprisingly, experiments on hBN-supported graphene devices still report a deviation from the linear increase of the resistivity with temperature, as expected for acoustic phonon scattering, in particular at lower densities near the CNP [10, 188]. This motivated Sohier and coworkers to revise the model of intrinsic phonon scattering in graphene [78]. The main results of their ab initio study are: i) acoustic phonon scattering is not caused by the deformation potential (which they find to

be strongly screened) but by the unscreened acoustic gauge field, but no significant correction to the above formula is needed, although both theories underestimate experimental values by $\approx 30\%$; ii) more importantly, electron-electron interaction renormalizes (increases) the coupling strength to *intrinsic optical phonon modes* at low densities, especially to the intervalley A_1' mode having the lowest energy and strongest coupling of all optical phonon branches. Although its activation energy is much higher than the thermal energy at room temperature, the substantially larger coupling strength still leads to a significant contribution to the resistivity. Hence, optical phonon scattering (by the A_1' mode) can become the dominant scattering mechanism near the CNP already at room temperature in devices with extremely low extrinsic disorder (see also Figure 2.9 and Section 2.5). The full intrinsic phonon model based on acoustic and optical branches, proposed by Sohier and coworkers, has been also adapted by other studies [80, 192].

2.3.3. Graphene at low density - a two carrier type system

An equal amount of electrons and holes always remains present at the CNP even though the gate controlled average density n_g is tuned to zero. As described in Section 2.2, they are either separated into electron and hole puddles by long range disorder in the "low" temperature limit (for $T < T_c$) or thermally excited transforming the spatially inhomogeneous landscape into a homogeneous system at high temperature. This residual total carrier density at the CNP leads to a finite minimum conductivity (also in the $T \to 0$ limit[40]). The Hall coefficient crosses smoothly through zero at the CNP instead of diverging as one would expect for a single carrier type system. In addition, a strong classical magneto-resistance arises. All these findings are a direct consequence of the coexistence of electrons and holes near the CNP.

In the following we will discuss three proposed models describing the electrical transport in the two carrier type regime (as well as its crossover to the single carrier type regime). The *two carrier type Drude model*, the extension of the Drude theory in order to account for two carrier types, is often employed in the literature. It assumes the system to be homogeneous and describes the system by (total) electron and

40 The disorder induced density fluctuations are spatially continuous. Because a band gap is absent, the boundaries between electron and hole puddles constitute smooth, highly transparent pn-junctions [3, 7, 31, 106, 193, 194]. Hence, electrical transport is not limited by scattering across these PNJ but by scattering within these puddles. This scattering can be described by (local) Boltzmann theory. As discussed in the previous sections, charge transport at the CNP is always diffusive, as a result of disorder, electron-hole or (optical) phonon scattering.

hole densities n_e, n_h, as already introduced in Section 2.2.4, and by the respective carrier mobilities μ_e, μ_h. The mobilities are ad-hoc parameters in the Drude model but can also be calculated in the Boltzmann theory for any given microscopic mechanism (see also Section 2.3.1). The assumption of a homogeneous system describes well the case when thermally excited carriers smooth the disorder induced spatial density inhomogeneity (e.g. for low disorder devices at high enough temperatures). Moreover, the two carrier type Drude model also serves as a fairly adequate approximation of the inhomogeneous system, since the puddle boundaries in graphene are quasi-transparent. Accordingly, in this work, we mostly use this model in combination with the channel approximation in order to obtain a good estimate of the residual carrier density, introduced as an ad-hoc parameter, and the mobility at the CNP from experimental data (in a phenomenological approach). We then refer to this approach as the two carrier (Drude) model within the channel approximation (see below). A more refined theory explicitly considering the spatial inhomogeneity is the *effective medium theory* (EMT). It is well applicable since the scattering across the puddle boundaries can be neglected. In its most general form, the model only depends on two parameters, the mobility μ (if electron and hole mobilities are equal) and the density fluctuations characterized by n_{rms} (see also Section 2.2.3). Both parameters can be obtained from fitting experimental data, or alternatively, an arbitrary microscopic theory, such as charged impurity disorder in the Boltzmann theory, can be implemented to calculate them. Finally, a third possibility is to account for the spatial inhomogeneity through a *modified DOS* of a homogeneous model system (see Section 2.2.4). The conductivity in this model system again can be calculated by the Boltzmann theory considering arbitrary microscopic scattering mechanisms.

Two carrier type Drude model within channel approximation (assuming spatial homogeneity)

When the electron and hole carrier density of the system can be considered as (spatially) homogeneous and electrical transport as diffusive, the two carrier type Drude model is applicable, e.g. in case of compensated semiconductors. It follows a phenomenological approach and describes the experimental results by ad-hoc (fitting) parameters which are a priori not connected to any microscopic theory. The expressions for the longitudinal resistivity and the Hall coefficient can be derived within the general Drude picture for coexisting electrons and holes (as shown in detail in

Appendix B.2). The conductivity in the absence of a magnetic field is then given by

$$\sigma_{xx}(B=0) = \frac{1}{\rho_{xx}(B=0)} = e\,(n_h\,\mu_h + n_e\,\mu_e) = e\,\mu\,n_t\,, \tag{2.56}$$

where in the last expression we assumed the electron and hole mobility to be equal, $\mu_h = \mu_e = \mu$ (and to be density independent). Using the already introduced *channel approximation* (see Equation 2.30), the total charge carrier density can be approximated very well by $n_t = \sqrt{n_{t,0}^2 + n_g^2}$ where $n_{t,0}$ is the (ad-hoc) residual total carrier density at the CNP. The average density $n_g = n_e - n_h$ is controlled by the applied gate voltage, $n_g = \beta\,(V_g - V_{CNP})$. The above equation can then be rewritten as follows

$$\sigma_{xx}(B=0) = e\,\mu\,\sqrt{n_{t,0}^2 + n_g^2} = e\,\mu\,\sqrt{n_{t,0}^2 + \beta^2\,(V_g - V_{CNP})^2}\,. \tag{2.57}$$

Containing four parameters ($n_{t,0}$, μ, β, and V_{CNP}), this equation can be used to fit the experimentally obtained $\rho(V_g)$ (or $\sigma(V_g)$) curve at zero magnetic field for a given temperature. Since all parameters may depend on temperature, the fitting has to be repeated for each temperature (resulting in $n_{t,0}(T)$ etc.). Typically, β is known from Hall measurements at higher densities, and V_{CNP} corresponds to the gate voltage of the resistivity maximum. Then, only the two remaining parameters, $n_{t,0}$ and μ, are adjusted in order to describe the $\sigma(n_g)$ curve.

A common and quick method[41] to estimate the residual carrier density at the CNP $n_{t,0}$ is based on the analysis of a double-logarithmic diagram of $\sigma(n_g)$, as shown for example in Refs. 18, 79, 80, 161, 196. $n_{t,0}$ then corresponds to the density value for which the extrapolation of the linear relation between $\log(\sigma)$ and $\log(n_g)$ (at higher densities) equals the minimum conductivity reached as a result of saturation at low carrier density[42]. Typically, this analysis is done manually by drawing lines in the double-logarithmic diagram (as demonstrated in panel a of Figure 3.8). If the residual carrier density is known, one can also use Equation 2.57 to derive a "two carrier" field effect mobility directly from the measured conductivity. We will use

41 An alternative to obtain $n_{t,0}$ from the measured conductivity curve is suggested in Refs. 126, 195. $n_{t,0}$ is determined by a direct fit of the $N(n_g) = \sigma/e\mu$ curve in a very narrow n_g range around the CNP ($n_g < n_{t,0}$).

42 This method implies that the mobility is (almost) density independent which should be a valid assumption in the density region of interest, since disorder scattering prevails at lower densities (leading to a constant mobility).

the notation $\mu_{\mathrm{FE},2}$ in analogy[43] to the single carrier type's definition given in Eq. 2.38 (see Section 2.3.1). It reads

$$\mu_{\mathrm{FE},2} = \frac{\sigma_{\mathrm{xx}}(\mathrm{B}=0)}{\mathrm{e}\sqrt{\mathrm{n}_{\mathrm{t},0}^2 + \mathrm{n}_{\mathrm{g}}^2}}\,. \tag{2.58}$$

In contrast to the field effect mobility definition in the single carrier type picture (diverging at the CNP) and the Hall mobility definition (zero at the CNP, see Eq. 2.42), the two carrier field effect mobility has a finite value at the CNP which we label μ_{CNP} (such that $\sigma_{\mathrm{min}} = \mathrm{e}\,\mathrm{n}_{\mathrm{t},0}\,\mu_{\mathrm{CNP}}$).

When a magnetic field is applied (perpendicular to the two-dimensional plane of graphene), the expressions for the longitudinal resistivity and the Hall coefficient in the two carrier type Drude theory become quite complex and bulky (given in Eq. B.18 and B.16 in Appendix B.2). Even in this general case, the channel approximation can be employed, since separate expressions for $\mathrm{n_e}$ and $\mathrm{n_h}$ can be derived, as shown in Eq. 2.31, and then inserted into the general two carrier type equations. Fortunately, several assumptions can be made in many cases which simplify the equations providing a direct understanding of the magnetic field dependence.

Typically, the electron and hole mobilities can be assumed to be equal, $\mu_{\mathrm{h}} = \mu_{\mathrm{e}} = \mu$. Then, the longitudinal resistivity is given by (see also Eq. B.19)

$$\rho_{\mathrm{xx}} = \frac{(\mathrm{n_h}+\mathrm{n_e})(1+\mu^2\mathrm{B}^2)}{\mathrm{e}\mu[(\mathrm{n_h}+\mathrm{n_e})^2 + (\mathrm{n_h}-\mathrm{n_e})^2\mu^2\mathrm{B}^2]}\,. \tag{2.59}$$

Exactly at the CNP ($\mathrm{n_e} = \mathrm{n_h}$) this simplifies to a quadratic magnetic field dependence $\rho_{\mathrm{xx}}(\mathrm{B}) = \rho_{\mathrm{xx}}(0)\,(1+\mu^2\mathrm{B}^2)$ and the classical magneto-resistance (MR) in the two carrier type Drude model reads

$$\mathrm{MR}_{\mathrm{two-carrier}} \equiv \frac{\rho_{\mathrm{xx}}(\mathrm{B}) - \rho_{\mathrm{xx}}(0)}{\rho_{\mathrm{xx}}(0)} = \mu^2_{\mathrm{MR,two-carrier}}\mathrm{B}^2\,. \tag{2.60}$$

We use here the subscript "MR,two-carrier" for the mobility since it serves as the only fitting parameter in the two carrier type Drude theory to describe the experimental MR data. If the two carrier type model describes perfectly the studied system, it should be equal to μ_{CNP}. In Section 4.5 we will compare the measured MR at the CNP with different models, including the two carrier type Drude theory. Finally, it is obvious that for classically weak magnetic fields ($\mu^2\mathrm{B}^2 \ll 1$) the MR becomes negligible and will vanish in the single carrier type regime (when $\mathrm{n_e}$ or $\mathrm{n_h} \to 0$).

43 As proposed by Ref. 197, one could also define $\mu_{\mathrm{LR},2}$ in analogy to Eq. 2.51 using $\rho_{\mathrm{xx}} = \rho_{\mathrm{LR}} + \rho_{\mathrm{SR}} = (\mathrm{e}\,\mu_{\mathrm{LR},2}\,\mathrm{n_t})^{-1} + \rho_{\mathrm{SR}}$.

For equal electron and hole mobility, the Hall coefficient is given by (Eq. B.17)

$$R_H = \frac{(n_h - n_e)\,(1 + \mu^2\,B^2)}{e\,[(n_h + n_e)^2 + (n_h - n_e)^2\,\mu^2\,B^2]}\,. \tag{2.61}$$

Obviously, R_H is zero exactly at the CNP ($n_e = n_h$) and approaches the single carrier type equation (Eq. 2.40) in the high density regime (when n_e or $n_h \to 0$). While R_H does not depend on the magnetic field in this regime, it exhibits a magnetic field dependence when $n_e \approx n_h$ (near the CNP) and when the magnetic field becomes so large that $\mu^2 B^2 \ll 1$ no longer holds. This $R_H(B)$ dependence is reflected in a non-linear magnetic-field dependence of ρ_{xy}, as shown in Figure 4.1 in Section 4.1. Hence, we restrict us in this work to classically weak magnetic fields (such that $\mu^2 B^2 \ll 1$). Then, R_H is constant (leading to linear Hall traces) and reads

$$R_H(\mu^2 B^2 \ll 1) = \frac{n_h - n_e}{e\,(n_h + n_e)^2} = -\frac{n_g}{e\,n_t^2} = -\frac{n_g}{e\,(n_{t,0}^2 + n_g^2)}\,. \tag{2.62}$$

Here, we used again the channel approximation in the last expression. This equation allows us to fit the measured Hall coefficient curve over the gate voltage (see also Ref. 82, 114, 142, 193, 198). Since $n_g = \beta\,(V_g - V_{CNP})$, in total three fitting parameters ($n_{t,0}$, β and V_{CNP}) can be adjusted, as discussed in more detail in Chapter 4. In contrast to the fitting routine for the longitudinal resistivity at zero magnetic field, the Hall coefficient does not depend on the mobility (in the two carrier type Drude model[44]). Therefore, the analysis of the Hall coefficient curves constitutes a valuable and complementary method to estimate $n_{t,0}$ which so far has not been addressed sufficiently in the literature. Moreover, from Equation 2.62 follows directly that the Hall coefficient extrema are $|R_{H,max}| = 1/2\,e\,n_{t,0}$ when $n_g = \pm n_{t,0}$.

The measured Hall coefficient curves sometimes exhibit an asymmetry between the electron and hole side (see also Section 4.2). This may originate from unequal electron and hole mobilities, characterized by $b = \mu_e/\mu_h$, and/or from different Hall scattering factors r_e, r_h, which are a result of the Boltzmann theory (see Section 2.3.1 and Equation 2.41) [150]. Equation 2.62 then transforms into

$$R_H(\mu^2 B^2 \ll 1) = \frac{n_h\,r_h\,\mu_h^2 - n_e\,r_e\,\mu_e^2}{e\,(n_h\,\mu_h + n_e\,\mu_e)^2} = \frac{n_h\,r_h - b^2\,n_e\,r_e}{e\,(n_h + b\,n_e)^2}\,. \tag{2.63}$$

The Hall coefficient becomes zero when the enumerator is zero, that means when $n_h\,r_h = b^2\,n_e\,r_e$. Hence, the zero-crossing of the Hall coefficient does not happen

44 In the effective medium theory accounting for the inhomogeneous landscape, the Hall coefficient near the CNP depends (very) slightly on the mobility. Typically, this can be neglected (private communication with Dr. Indra Yudhistira, theoretician in the group of Prof. Shaffique Adam at NUS, co-authors of Ref. 110).

exactly at $V_g = V_{CNP}$ (when $n_e = n_h$ or $n_g = 0$)[45]. Using the separate expressions for n_e, n_h in the channel approximation (Eq. 2.31), Equation 2.63 can be rewritten as follows (still restricting us to $\mu^2 B^2 \ll 1$)

$$R_H = \frac{2}{e} \frac{(r_h - r_e\, b^2)\sqrt{n_{t,0}^2 + \beta^2 (V_g - V_{CNP})^2} - (r_h + r_e\, b^2)\, \beta\, (V_g - V_{CNP})}{\left[(b+1)\sqrt{n_{t,0}^2 + \beta^2 (V_g - V_{CNP})^2} + (b-1)\, \beta\, (V_g - V_{CNP})\right]^2}\,. \quad (2.64)$$

One can show that all parameters cannot be used at once for the curve fitting. For instance, changing β by an arbitrary factor m yields the exact same solution when $n_{t,0}$ and r_e, r_h scale with the same factor m. This does not lead to a unique minimum in the fitting procedure. Therefore, certain parameters such as β and r_e, r_h are obtained from the single carrier type regime at high densities, keeping them then fixed when fitting the whole $R_H(V_g)$ curve (see Section 4.2).

Effective medium theory accounting for spatial inhomogeneity

At first, we briefly summarize the previous sections. The two carrier type Drude theory assumes the system to be spatially homogeneous and describes the transport measurements phenomenologically. The residual carrier density and the mobility are ad-hoc fitting parameters and not (necessarily) connected to a microscopic theory. The Boltzmann theory, presented in Section 2.3.1, requires the introduction of a microscopic scattering mechanism (for concreteness). It is restricted to a homogeneous single carrier type system leading to $\sigma \to 0$ for $n_g \to 0$ in the low temperature limit (as shown in panel a of Figure 2.7). Since this disagrees with the experimental findings, a self-consistent approximation model (SCA) was proposed (see Sections 2.2.2 and 2.3.2). It also requires the consideration of a microscopic theory for concreteness. At low density ($n < n_*^{SCA}$), the minimum conductivity plateau is then approximated by a homogeneous model system with residual density n_*^{SCA} such that the experimentally obtained σ_{min} equals the Boltzmann conductivity $\sigma_B(n_*^{SCA})$ in the low temperature limit [7, 29, 99]. Panel a of Figure 2.7 shows the "hard" plateau as a result of the SCA model and its immediate, discrete crossover to the Boltzmann conductivity $\sigma_B(n)$ (for $n > n_*^{SCA}$). Although it constitutes a reasonable approximation, this discrete crossover is artificial. Moreover, the SCA model is limited conceptually to the $T \to 0$ case.

45 Strictly speaking, V_{CNP} ($n_e = n_h$) also does not correspond to the resistivity maximum, given by $\min(\mu_e\, n_e + \mu_h\, n_h)$, if the electron and hole mobilities are drastically different which is however typically not the case in graphene.

In contrast to all these theories, the effective medium theory (EMT) generally describes the electrical transport through a heterogeneous system with spatial inhomogeneities (typically in terms of fluctuating density) [106]. In its general form, it does not require knowledge about the origin of these spatial inhomogeneities. However, the theory is restricted to the case when the total conductivity of the inhomogeneous system is governed by the conductivity within the patches and not between them [7, 106]. This condition is fulfilled for graphene at the CNP, since the pn-junctions between the puddles constitute quasi-transparent boundaries due to the absence of a band gap [7, 106]. As shown in Refs. 7, 29, 106, the system's total conductivity in the effective medium theory σ_{EMT} for a gate controlled average carrier density $\mathrm{n_g}$ can be found by solving the following equation numerically

$$\int_{-\infty}^{\infty} \mathrm{dn}\ \mathrm{P}[\mathrm{n}, \mathrm{n_g}, \mathrm{n_{rms}}]\, \frac{\sigma(\mathrm{n}) - \sigma_{\mathrm{EMT}}}{\sigma_{\mathrm{B}}(\mathrm{n}) + \sigma_{\mathrm{EMT}}} = 0\,. \tag{2.65}$$

$\sigma(\mathrm{n})$ is the local conductivity (of a small homogeneous patch with density n), $\mathrm{P}[\mathrm{n}, \mathrm{n_g}, \mathrm{n_{rms}}]$ is the probability distribution of the carrier density in the inhomogeneous system, and $\mathrm{n_{rms}}$ characterizes the root mean square density fluctuations. In order to obtain analytical expressions, typically a Gaussian[46] distribution $\mathrm{P^G}[..]$ is assumed [7, 29, 106, 109]. Then, at low temperature, it follows that $\mathrm{n_{rms}} \approx \sqrt{3}\mathrm{n_*^{SCA}}$ (see also Eq. 2.27), and accordingly $\sigma_{\mathrm{min}}(\mathrm{T} \to 0) = \mathrm{n_{rms}}\, \mathrm{e}\, \mu/\sqrt{3}$ [7, 29, 110]. In a general, phenomenological approach, the local conductivity $\sigma(\mathrm{n})$ in Eq. 2.65 can be determined by simply introducing an ad-hoc (density-independent) mobility, such that $\sigma(\mathrm{n}) = \mathrm{e\,n}\,\mu$. In doing so, $\mathrm{n_{rms}}$ and μ serve as independent fitting parameters in the EMT model to describe the measured density dependence of the conductivity at low temperature [110].

A better approach is to incorporate a microscopic disorder and scattering model by using the Boltzmann theory to obtain the local conductivity $\sigma_{\mathrm{B}}(\mathrm{n})$ (in Eq. 2.65). So far, this has been done only for charged impurity disorder [7, 29, 106, 109, 110]. As shown in the previous sections, charged impurity disorder leads to a density-independent mobility μ_{c} and a residual carrier density at the CNP $\mathrm{n_{*,c}^{SCA}}$. Both are determined self-consistently through the three modeling parameters of the charged impurity disorder model $(\mathrm{n_{imp}}, \mathrm{d}, \mathrm{r_s})$. Hence, both EMT parameters, $\mathrm{n_{rms}}$ $(= \sqrt{3}\,\mathrm{n_{*,c}^{SCA}})$ and the local Boltzmann conductivity $\sigma_{\mathrm{c}}[\mathrm{n_{imp}}, \mathrm{d}, \mathrm{r_s}]$, are well defined. Panel a of Figure 2.7 nicely demonstrates that the EMT model for charged impurity

46 Alternatively, also a Lorentzian distribution can be assumed, or P[..] can be computed using the microscopic Thomas Fermi Dirac theory as performed in Refs. 29, 106. The TFD approach describes best the crossover region between the plateau and the high density regime in graphene and accounts for many body effects. However, the use of the Gaussian distribution results in a fairly good approximation of the TFD approach [29].

disorder connects smoothly the minimum conductivity at the CNP with the Boltzmann limit at high density (in the low temperature limit). Finally, we emphasize that the effective medium theory (Eq. 2.65) would be also applicable if the mobility was density-dependent and/or explicitly temperature dependent (that means when also other scattering mechanisms such as phonon scattering were included in the local Boltzmann theory)[47].

Apart from describing the density dependence of the conductivity in the low temperature limit, the effective medium theory also allows modeling the temperature dependence of the conductivity (as long as the disorder induced density inhomogeneity is not smeared substantially by thermally activated carriers). When the temperature is increased, the local conductivity is affected by the thermal activation of charge carriers as well as by the temperature dependence of the scattering mechanisms[48]. Eventually, at higher temperatures, the thermally activated carriers will smoothen entirely the disorder induced electron hole puddle landscape. Then, the EMT theory with the parameter n_{rms}, characterizing the disorder induced density fluctuations in the low temperature limit, is no longer a proper approach to describe this homogeneous system at the CNP. As already discussed in Section 2.2.3, the disorder equivalent temperature scale is defined as $T_* = E_F(n_{rms})/k_B$ (see also Eq. 2.28). The critical temperature, when the thermally excited carrier density equals the disorder induced puddle density, is given by $T_c \approx 0.55\, T_*$ (see Ref. 80 and Eq. 2.29). For our best hBN-encapsulated devices with $n_{t,0}(T \to 0) \approx 10^{10}\,cm^{-2}$, this critical temperature is in the order of 100 K, while it can become even larger than room temperature for bare devices on SiO_2 with higher disorder level (as shown in Table 2.1). Below T_c, the temperature dependence of the conductivity can be described by the effective medium theory considering a temperature dependent local conductivity but fixed n_{rms} value (obtained at low temperature).

So far, all theoretical studies discussing the temperature dependence of the EMT conductivity focused on charged impurity disorder as main scattering mechanism and source of the density fluctuations at the CNP [7, 80, 109]. For simplicity, the weak explicit temperature dependence of the Coulomb scattering, due to the temperature dependent screening function, is neglected. Hence, the mobility μ_c is assumed

47 However, this has not yet been addressed in the published studies. Only Ref. 110 introduces an additional "short-range", density-independent conductivity σ_s which is employed as additional fitting parameter at low temperature, such that $\sigma_B(n_{imp}, d, r_s, \sigma_s)$.

48 This arises both from the thermal energy averaging in the Boltzmann theory (Eq. 2.44) as well as from the explicit temperature dependence of the scattering mechanism, such as the temperature dependent dielectric screening function for charged impurity scattering (as already discussed in Section 2.3.2) [160].

to be temperature independent. Together with n_{rms}, it is determined[49] at low temperature, and then both are used as fixed parameters. In doing so, the only contribution to the temperature dependence of the local conductivity originates from the thermal activation of charge carriers (the thermal smearing of the Fermi surface) [7, 109]. The local conductivity can be written as $\sigma(n, T) = e\, n_{rms}\, \mu_c\, H(n/n_{rms}, T/T_*)$. H is a dimensionless function that describes the thermal activation of electrons and hole (using n_e, n_h as given in Eq. 2.22 in Section 2.2.2) [7, 109]. Therefore, the temperature dependence of the local conductivity, and so the temperature dependence of the EMT conductivity, is governed solely by the density of states in graphene (for a scattering mechanism such as charged impurities that to good approximation is not explicitly temperature dependent [160]). Hence, it is determined by universal properties, such as the type of band dispersion (linear for monolayer or approximately parabolic for bilayer graphene) and its parameters (Fermi velocity v_F for monolayer or effective mass m^* for bilayer).

It was then shown in Refs. 7, 80, 109 that, below T_c, the scaled EMT conductivity $\sigma_{EMT}/e\, n_{rms}\, \mu_c$ is a universal function of the scaled temperature T/T_* and the scaled average density n_g/n_{rms} (universal scaling, as shown in panel b and c of Figure 2.7). For each experimentally studied device, the EMT parameters μ_c and n_{rms} are obtained from the low temperature transport measurements, following the aforementioned procedure. They are then used to scale the temperature, the density, and the measured conductivity. Typically, the universal one-parameter scaling (for charged impurities leading to a temperature-independent mobility) can be demonstrated by studying the dependence of the scaled minimum conductivity at the CNP $\sigma_{min}(T)/e\, n_{rms}\, \mu_c$ on T/T_*. In doing so, the data points obtained on several bilayer devices on SiO_2 with different disorder level collapse on the theoretically predicted universal curve (see panel c of Figure 2.7) [7, 109]. Alternatively, it is also illustrative to study the broadening of the conductivity minimum. The latter is characterized by the residual carrier density $n_{t,0}(T)$ which is obtained, for each temperature, from the double-logarithmic diagram of the measured density dependence of the conductivity and then scaled by $n_{rms} = \sqrt{3} n_{t,0}(T \to 0)$. The same procedure can also be performed on the calculated $\sigma_{EMT}(n_g/n_{rms})$ curves (for each temperature increment). Experimental measurements of several hBN-encapsulated

49 The low temperature values of n_{rms} and μ_c are either calculated using the model parameters of the charged impurity disorder (n_{imp}, d, r_s) or, as typically done in practice, they are obtained directly from low temperature (transport) experiments. n_{rms} either can be determined by scanning local probe experiments or it can be estimated from the residual total density at the CNP ($n_{rms} \approx \sqrt{3} n_{t,0}(T \to 0)$). This residual density is estimated in a double-logarithmic diagram of the measured $\sigma_{exp}(n_g)$ curve, as discussed earlier [80]. The low temperature mobility μ_c is derived either from the linear density dependence of the measured conductivity at higher densities, or simply via $\sigma_{min,exp}(T \to 0) = e\, n_{rms}\, \mu_c/\sqrt{3}$.

monolayer graphene devices and the universal theoretical prediction for $n_{t,0}(T)/n_{rms}$ are compared in the inset of Figure 2.7c [80]. Again, only for temperatures above T_c, the data points differ from the universal curve which the authors of Ref. 80 attribute to the smoothening of the density inhomogeneity above that temperature and, as a consequence, to the failing of the effective medium theory. Another plausible argument is that additional scattering mechanisms have to be considered and that the Fermi velocity might be larger than the assumed value of $v_F = 1 \cdot 10^6\,m/s$ in the theory[50]. In Section 4.3 we will discuss in more detail the temperature dependence in this regime including the effect of a renormalization of the Fermi velocity.

In the presence of a magnetic field, the magneto-transport is described in the effective medium theory by two coupled equations which read [110]

$$\begin{aligned} \int_{-\infty}^{\infty} dn\ P[n, n_g, n_{rms}]\ \frac{\sigma_{xx}^2[n] - (\sigma_{xx}^{EMT})^2 + (\sigma_{xy}^{EMT} - \sigma_{xy}[n])^2}{(\sigma_{xx}[n] + \sigma_{xx}^{EMT})^2 + (\sigma_{xy}^{EMT} - \sigma_{xy}[n])^2} &= 0 \\ \int_{-\infty}^{\infty} dn\ P[n, n_g, n_{rms}]\ \frac{\sigma_{xy}[n] - \sigma_{xy}^{EMT}}{(\sigma_{xx}[n] + \sigma_{xx}^{EMT})^2 + (\sigma_{xy}^{EMT} - \sigma_{xy}[n])^2} &= 0\,. \end{aligned} \tag{2.66}$$

Here, $\sigma_{xx}[n], \sigma_{xy}[n]$ are again the local conductivities in a homogeneous patch of the system, while $\sigma_{xx}^{EMT}, \sigma_{xy}^{EMT}$ are the total conductivities of the inhomogeneous system at a given gate controlled average density n_g. With these conductivities one can then calculate the magneto-resistance or Hall coefficient in the effective medium theory. So far, this has been done only for the magneto-resistance at low temperature in Ref. 110 assuming charged impurity disorder (that means a density independent mobility). For classically weak magnetic fields ($\mu^2 B^2 \ll 1$), the magneto-resistance can be described by

$$MR_{EMT}(n_g, B) = A[n_g/n_{rms}]\ \mu_{MR,EMT}^2\ B^2\,. \tag{2.67}$$

The dimensionless pre-factor $A[n_g/n_{rms}]$ depends universally on the scaled density, as demonstrated in Ref. 110 for several devices with different disorder level. At the CNP, one obtains $A[0] = 1/2$. Then, it follows that $\mu_{MR,EMT} = \sqrt{2}\mu_{MR,two-carrier}$ when both mobilities are used as fitting parameters to describe the same measured MR curve at the CNP (compare to Eq. 2.60 and see also Section 4.5)[51]. However, the effective medium theory describes quantitatively better the amplitude of the

50 Private communication with Prof. Shaffique Adam and Dr. Indra Yudhistira, authors of Ref. 80.

51 Here, we have considered already the advanced EMT model which accounts for the continuous distribution of electron hole puddles. Earlier studies were based on an area-fraction EMT model which predicts $MR_{areaEMT} = \sqrt{1 + \mu_{areaEMT}^2\,B^2} - 1$ at the CNP (see also Section 4.5) [158, 199]. For classically weak magnetic fields, the MR curve at the CNP is almost identical for both EMT models, the area-fraction and the continuous one.

measured MR in the crossover regime to the single carrier type case (as a function of n_g/n_{rms} and when assuming the mobility to be density independent) [110]. In comparison to the two carrier type Drude model, we also expect that the EMT model will allow us to quantitatively describe the measured Hall coefficient curve as a function of gate voltage or average density better (and even as a function of the renormalized n_g/n_{rms}). In particular, in the crossover region near the CNP when $n_g/n_{rms} \approx 1$, the Hall coefficients should (slightly) differ including the Hall coefficient extrema. The deviations are more pronounced at low temperature and/or for devices with high disorder level, where the system exhibits a larger degree of spatial inhomogeneity near the CNP. Using the EMT model to fit the measured Hall coefficient curves (at low temperature) also allows us to obtain directly n_{rms} and the density independent mobility μ. This serves as an important tool complementing the approach that fits the conductivity versus density at zero magnetic field, since the Hall coefficient curve is more sensitive to n_{rms} but less sensitive to the mobility (see Chapter 4 for further discussion).

Coexistence of diffusive and activated transport - the modified DOS approach

The EMT model fails to explain quantitatively the temperature dependence of the conductivity above T_c (at least when phonon scattering and renormalization of the band dispersion are not considered). Conceptually, it has the major drawback that it only considers the temperature dependence of the local conductivity, i.e. the contribution of the diffusive transport inside the puddles. The group of Das Sarma has therefore developed a two-component transport model which accounts for the coexistence of diffusive transport inside the puddles and activated transport across the disorder induced potential fluctuations [97, 111, 115, 149]. This approach then explains why even away from the CNP an insulating (increasing conductivity with temperature) or at least a non-monotonic behavior instead of the expected purely metallic behavior (decreasing conductivity) can be observed in devices with high disorder level [97, 111]. For devices with low disorder, the insulating behavior is observed within a more narrow density range around the CNP but it is then more pronounced than for devices with high disorder. Eventually, devices with extremely low disorder level exhibit a very low minimum conductivity at the CNP in the low temperature limit. In other words, more ordered graphene devices (with higher quality) manifest a more disordered (stronger insulating) behavior [149]. This can be understood using the two-component transport model which is introduced briefly in the following and in detail in Refs. 97, 111.

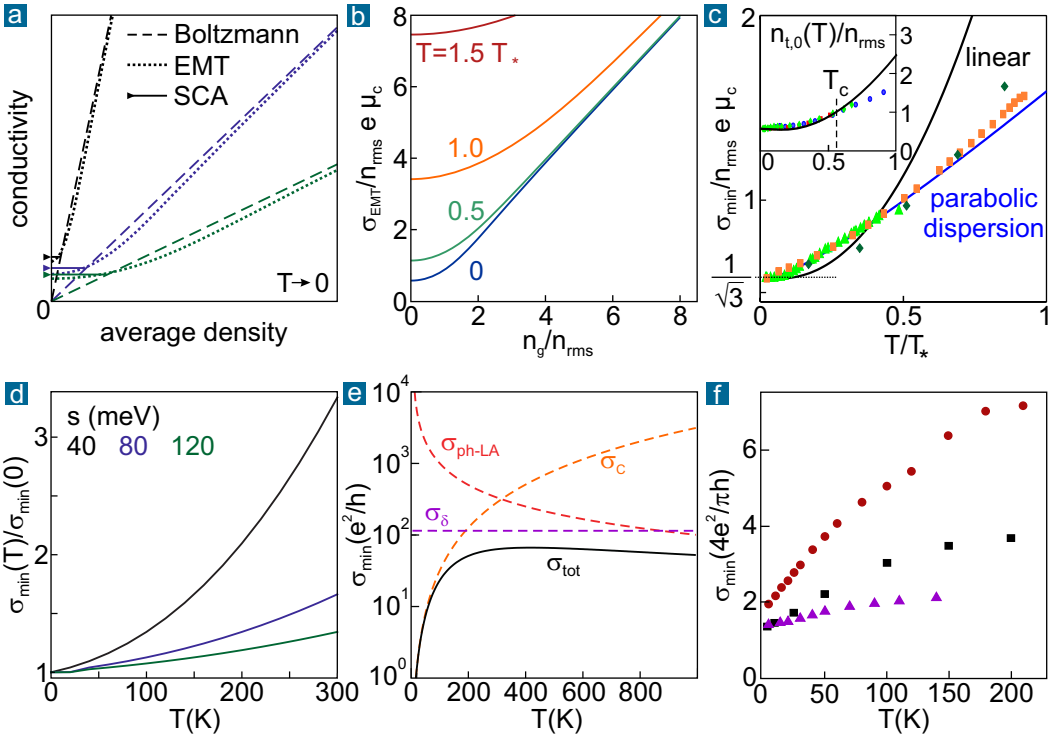

Figure 2.7. Theoretical models describing the density and temperature dependence of the conductivity. a: Density dependence of the charged impurity disorder limited conductivity for $T \to 0$, comparing the single carrier type Boltzmann theory, the SCA theory, and the EMT (modified from Ref. 7). The impurity density n_{imp} is varied from low (black) to high (green) values. b: Scaled EMT conductivity as a function of scaled density for various temperatures (modified from Ref. 7). c: For charged impurity disorder, the EMT model finds the scaled minimum conductivity to follow a universal function of the scaled temperature, see also main text (modified from Ref. 109). The main figure compares bilayer (parabolic) and monolayer (linear dispersion) graphene and several bilayer devices (data points). The inset shows the universal scaling behavior of the residual carrier density at the CNP for monolayer graphene (from Ref. 80). d: Temperature dependence of the minimum conductivity based on the modified DOS model (Eq. 2.70) for different disorder strength parameters (from Ref. 97). Phonon scattering is not considered. e: Temperature dependence of σ_{min} in the Boltzmann theory assuming the absence of a disorder induced carrier density at low temperature but taking into account several scattering mechanisms (from Ref. 160). f: Measured two-terminal minimum conductivity of three suspended graphene devices (from Ref. 126). The lower the disorder level (red dots), the more pronounced is the insulating behavior. The shape of the curve can be explained qualitatively by combining panels d and e.

Referring to Section 2.2.4, the density fluctuations are assumed to follow a Gaussian distribution and are phenomenologically described by a modified DOS (see Eq.

2.33). Apart from the Fermi velocity v_F, a legacy of graphene's original DOS, the modified DOS is controlled by a single parameter, the standard deviation $s \equiv V_{rms}$ of the Gaussian potential fluctuations[52]. The average conductivity of the electron and hole puddles is then given by $\sigma_{e/h} \propto n_{e/h} \langle\tau\rangle$. Here, $n_{e/h}$ are the average electron and hole densities obtained from Equation 2.34 and $\langle\tau\rangle$ is the average transport relaxation time calculated within the Boltzmann transport theory for all considered scattering mechanisms[53]. The activated conductivities can be described by $\sigma^{(a)}_{e/h}(V) = \sigma_{e,h} \exp[\pm(E_F - V)/k_BT]$. The model then sums up the electron (hole) puddles as region 1 (2) occupying the area fraction $p = \int_{-\infty}^{E_F} P(V)dV$ and $1-p$, respectively, such that [97]

$$\begin{aligned} \sigma_1 &= \frac{1}{p} \int_{-\infty}^{E_F} (\sigma_e + \sigma_h^{(a)})\, P(V)\, dV \\ \sigma_2 &= \frac{1}{1-p} \int_{E_F}^{\infty} (\sigma_e^{(a)} + \sigma_h)\, P(V)\, dV\,. \end{aligned} \tag{2.68}$$

Finally, the total conductivity σ_t is given as a result of a 2D binary mixture of these two components and reads [97]

$$\sigma_t = (p - 0.5) \left[(\sigma_1 - \sigma_2) + \sqrt{(\sigma_1 - \sigma_2)^2 + \frac{4\,\sigma_1\,\sigma_2}{(2p-1)^2}} \right]. \tag{2.69}$$

The competition between the three energies (E_F, s, and k_BT) makes the temperature dependence of the conductivity at finite, moderate doping quite complex and non-monotonic, as discussed in Ref. 97. However, at the CNP ($E_F = 0$) and for $k_BT \ll s$, the minimum (total) conductivity reduces to [97]

$$\sigma_{min}(T)/\sigma_{min}(0) = 1 + \sqrt{\frac{2}{\pi}}\frac{k_BT}{s} + \frac{\pi^2}{3}\left(\frac{k_BT}{s}\right)^2. \tag{2.70}$$

Panel d of Figure 2.7 shows the result of Eq. 2.70 for three different disorder parameter values s. Note that phonon scattering is not considered which would otherwise lead to a saturation of the conductivity and then to a decrease at high temperature, as shown in panel e. The lower the disorder level (the smaller s), the more pronounced is the insulating behavior at the CNP. For most graphene devices, i.e.

52 The disorder or inhomogeneity parameter s is typically used as free fit parameter to describe experimental data. It is not necessarily linked to a microscopic theory (although this could be done, e.g. for charged impurity disorder, since $s^2 \propto n_{rms}$ [149]).

53 For instance, the theoretical study of Das Sarma and coworkers (Ref. 97) considers charged impurities as long range and zero-range potentials as short range disorder but does not include phonon scattering. Alternatively, to good approximation, one could also assume a density and temperature independent mobility μ as fitting parameter, such that $\sigma_{e/h} = e\, n_{e/h}\, \mu$.

statistically for the majority of the hBN-encapsulated graphene devices and for all uncovered graphene samples on bare SiO_2, the disorder parameter s is close to or even (much) larger than the thermal energy at room temperature. This explains why early transport measurements conducted on low quality devices on SiO_2 found the minimum conductivity to be almost temperature independent within the typical temperature range of these experiments (from low to room temperature) [149].

Eventually, when $s \to 0$ and/or $k_B T \gg s$, the two-component transport model approaches the Boltzmann transport theory. Hence, for low but finite temperature, the minimum conductivity at the CNP exhibits a power-law temperature dependence $\propto T^{\alpha}$ [149, 160]. The exponent α varies between 0 and 2, depending on the dominant scattering mechanism [160]. For long range scattering mechanisms that possess only a weak explicit temperature dependence, such as charged impurity disorder, the temperature dependence is quadratic ($\alpha = 2$). This power-law temperature dependence of the minimum conductivity and the disorder-by-order phenomenon are a direct consequence of the linear and gapless band structure of graphene ($\sigma_{min} \propto n_{t,0} \propto T^2$ follows from Eq. 2.35) [149, 160]. Moreover, this behavior can be easily distinguished from the exponential temperature dependence related to Anderson localization or gap-induced insulating behavior (for instance, when graphene is well aligned on hBN, see Section 2.1.6). Since $\sigma_{min}(T \to 0) \propto s^2$, the resistivity at the CNP may reach high values in the low temperature limit in case of devices with extremely low disorder ($s \to 0$). However, the temperature dependence will always follow a power-law and $\sigma_{min}(T \to 0)$ will always saturate towards a finite, disorder level dependent value at low enough temperature, $k_B T \ll s$ (empirically $k_B T_{sat} \approx s/5$) [149, 160]. For the cleanest suspended graphene devices with a disorder induced residual carrier density at the CNP of only about $10^8\,cm^{-2}$, the minimum conductivity starts to saturate below 10 K [126, 160]. In order to access intrinsic Dirac point physics at very low temperature ($\leq 100\,mK$), such as the Anderson localization threshold, the disorder induced density inhomogeneities should be less than $10^4\,cm^{-2}$ which is likely impossible to achieve in real devices [160]. Therefore, we observe the here discussed phenomena, the saturation of $\sigma_{min}(T \to 0)$ and a power-law temperature dependence, in all our devices (see Section 4.3).

The two-component transport model has not yet been developed further to describe the transport phenomena in the presence of a magnetic field. However, as already discussed in Section 2.2.4, the average electron and hole densities calculated from the modified DOS (Eq. 2.34) can be used in the equation for the Hall coefficient in the two carrier type Drude theory (Eq. 2.62). Then, it can be demonstrated that the channel approximation agrees well with the modified DOS model when the disorder parameter s and the residual carrier density at the CNP $n_{t,0}(T \to 0)$

are chosen appropriately (see panel f of Figure 2.5)[54]. Hence, we will employ the two carrier type Drude model within the channel approximation to analyze our magneto-transport measurements in this work.

2.3.4. Summary of the temperature dependence of diffusive transport

Graphene possesses a density tunable "metal-insulator transition" (MIT) when defining metallic behavior as $\partial\sigma/\partial T < 0$ and insulating behavior as $\partial\sigma/\partial T > 0$. The semiclassical theory for diffusive transport includes five temperature-dependent contributions as summarized in Ref. 149: 1) thermal activation of electron-hole occupancy; 2) energy averaging in the Boltzmann conductivity (Eq. 2.44); 3) thermal activation of carriers over disorder induced potential fluctuations; 4) indirect temperature dependence of the scattering mechanism, e.g. due to the screening function $\epsilon(q, T)$; 5) direct temperature dependence of the scattering mechanism, typically phonon scattering. In general, increasing the average density (or E_F) will suppress all these contributions.

In the single carrier type regime at high density (when $E_F \gg k_B T, T_*$), only the last two contributions play a role and the system shows metallic behavior $\partial\sigma/\partial T < 0$. It is mainly governed by (acoustic) phonon scattering. Note that we assume here the system to be large enough so that transport is always diffusive (and does not enter the ballistic regime). At the CNP, all above listed effects have to be taken into account and the temperature dependence of the conductivity is non-universal and sample dependent (in terms of the disorder level and the dominant source of scattering), as shown in panel f of Figure 2.7 and in Section 4.3. A more pronounced insulating behavior at the CNP arises for devices with lower disorder, while such low disorder will at the same time result in higher mobility (and thus larger conductivity or lower resistivity) at high density. Therefore, the density-tunable MIT is more distinct in cleaner devices with extremely small density fluctuations, which was for example achieved experimentally by shielding remote charged impurities[55]. For such devices, the minimum conductivity approaches a disorder level dependent

54 Moreover, this allows us to determine the value of the disorder parameter from the residual carrier density at the CNP, $n_{t,0}(T \to 0) = s^2 (\pi\hbar^2 v_F^2)^{-1}$. The residual density can be easily estimated using the double-logarithmic diagram method.

55 In addition to the here discussed graphene-intrinsic effects seen for suspended devices, such as the power-law insulating behavior and the saturation of the minimum conductivity at low enough temperature, the special arrangement of graphene-hBN heterostructures consisting of two graphenic sheets or with an additional top gate in close distance can cause another insulating behavior at the CNP in the limit of very low temperature [149, 200–202]. However, our devices with conventional design will not be affected by such effects.

value in the low temperature limit and follows a power-law temperature dependence which is peculiar to graphene's gapless linear band structure. Eventually, phonon scattering reverses this insulating behavior and the conductivity starts to decrease again at higher temperatures (metallic behavior). Since phonon scattering is weak in graphene, the onset of the trend reversal happens around room temperature which constitutes the upper temperature limit for most transport measurements. Thus, the minimum conductivity not only saturates at low temperature but also appears to saturate around room temperature. Moreover, the minimum conductivity at the CNP for high disorder devices exhibits almost no temperature dependence in the typically studied window up to room temperature, since $k_BT \ll s$. Finally, we emphasize that the conductivity (and its temperature dependence) at the CNP is entirely described by diffusive transport, since the disorder level is high enough in all realistic devices such that the intrinsic Dirac point physics in the low temperature limit is obscured. We also note that the temperature dependence of the conductivity in the crossover region (between the CNP and the single carrier type regime) is even more complex and governed by the interplay between three energies (E_F, s, and k_BT).

In this work, we focus on magneto-transport measurements (for classically weak magnetic fields). This gives us the opportunity to study the temperature dependence of the Hall coefficient curves together with the temperature dependence of the conductivity at zero magnetic field. Since the Hall coefficient is mostly sensitive to the carrier densities (including their spatial fluctuations), and less sensitive to the scattering processes, it constitutes a complementary tool to study directly the thermal activation of carriers in graphene, that means in a system with disorder induced spatial inhomogeneity at low temperature.

2.4. Mesoscopic transport

In the previous section we have assumed the momentum relaxing rate to be the largest in the system, and the dominant momentum relaxing mechanism to be carrier-impurity and carrier-phonon scattering. However, when sample quality is improved by reducing impurity density and when the temperature is lowered, the diffusive mean free path $L_{c-imp/ph}$ can exceed the device dimensions (length L and width W) and the mesoscopic transport regime is entered. This regime further splits into two cases for graphene. In the first and more familiar case known from other materials, carrier-carrier interactions play a minor role and their characteristic length scale L_{c-c} is much larger than the device dimensions ($L \ll L_{c-imp/ph}, L_{c-c}$). Then,

the only scattering mechanism causing momentum relaxation and affecting the trajectories of the non-interacting charge carriers occurs at the boundaries of the sample. This corresponds to the purely ballistic transport regime. It is realized at low temperature and for higher densities.
In the second case (related to section 2.1.7) the carrier-carrier interaction related length scale L_{c-c} can become the smallest in the system ($L_{c-c} \ll L_{c-imp/ph}, L$) for intermediate temperatures and smaller densities (see also Figure 2.9a). Then momentum conserving (MC) collisions dominate over all other momentum relaxing (MR) mechanisms (boundary scattering and diffusive scattering). Carrier transport becomes collective and will exhibit viscous features. The system can then be described by hydrodynamic quantities, as in classical fluid dynamics.

2.4.1. Ballistic transport

While charge transport near the CNP is always diffusive over the whole temperature range, at high enough densities ballistic transport can exceed several tens of micrometer in suspended and hBN supported graphene at low temperature [10, 196, 203, 204]. It can even reach the micrometer scale at room temperature [10, 195]. Ballistic features are more pronounced in narrow channels. The conductivity through a constriction with width W is then described in the low temperature limit by the Sharvin formula [68, 204, 205]

$$\sigma_{ball} = \frac{4e^2}{\pi h} W \sqrt{\pi n}\,. \tag{2.71}$$

This square root dependence on the density can be easily confused with the interplay of diffusive long range and short range scattering (see Figure 2.6c) and, therefore, the ballistic limit should always considered when analyzing low temperature data. A direct proof of ballistic transport is the observation of a negative bend resistance $R_B(n)$ for densities above a certain crossover density $n_c(T)$ [10, 81, 195, 203, 205]. Furthermore, graphene has been found to be a promising system to investigate electron optics, since its carriers move unimpeded over long distances and p-n junctions are highly transparent [206–208].

One important question for our work is how magnetotransport quantities are changed in the ballistic regime. This question was addressed recently in a general theoretical work by Scaffidi and coworkers [209]. It also describes qualitatively the behavior in graphene, although their theory does not explicitly account for the chiral character of its carriers which modifies boundary scattering [162]. In the fully ballistic case ($W < L_{MR}$ and $L_{MC}/L_{MR} = 10$), the longitudinal and the transverse (or Hall)

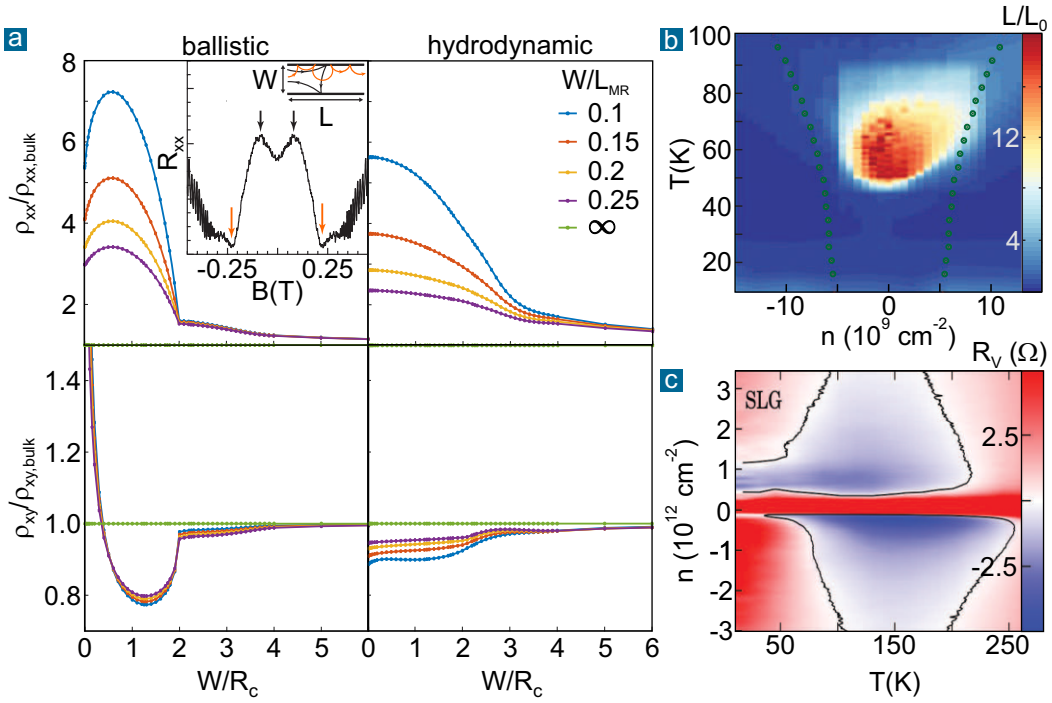

Figure 2.8. Ballistic and hydrodynamic transport in graphene. a: Theoretical prediction for the magneto-resistance and Hall resistivity in the ballistic ($L_{MC}/L_{MR} = 10$) and hydrodynamic case ($L_{MC}/L_{MR} = 0.05$) depending on W/R_c, where W is the channel width and $R_c = \hbar\sqrt{\pi n}/eB$ the cyclotron radius in monolayer graphene (taken from Ref. 209). Inset shows measured MR in ballistic constrictions peaking at $W/R_c = 0.9 \pm 0.1$ (from Ref. 162). b: Violation of Wiedemann-Franz law at intermediate temperature at the CNP in graphene with extremely low extrinsic disorder. Measured Lorenz ratio exceeds the expected ratio up to a factor of 20 indicating the emergence of a charge neutral hydrodynamic regime (from Ref. 79). c: Negative vicinity resistance R_V can be observed in high quality graphene hall-bars due to electron backflow (so-called whirlpools) (from Ref. 210). This can be explained within the hydrodynamic theory when electron-electron collision rate dominates and the electron liquid becomes viscous. Note the excellent qualitative agreement to the theoretical prediction of the hydrodynamic window shown in Figure 2.9.

resistivity differ from their corresponding bulk values in the diffusive regime (defined by $W/L_{MR} \to \infty$), here given as a function of the ratio between the channel width W and cyclotron radius R_c (see Figure 2.8a). While to our knowledge ballistic Hall resistivity measurements have not yet been reported, Masubuchi and coworkers [162] have observed characteristic magneto-resistance peak features at finite density (away from CNP) in two-terminal graphene on hBN constrictions where in the semi-classical single carrier type picture no magneto-resistance should emerge (inset of Figure 2.8a). They find the peaks to be located at $W/R_c = 0.9 \pm 0.1$ in

contrast to the general theory for conventional 2DEG which predicts $W/R_c = 0.55$. This suggests that boundary scattering is altered by the chiral nature of graphene's charge carriers.

2.4.2. Hydrodynamic behavior and viscous transport

In very clean devices of at least some few micrometer in size, the carrier-carrier collision related length L_{c-c} (see Equation 2.15) can become the smallest length scale of the system. Figure 2.9a shows a calculation of the ratio between the carrier-carrier scattering rate and the largest momentum-relaxing scattering rate as a function of temperature and density (for a system with low disorder). In the case of monolayer graphene, the hydrodynamic window has a characteristic lung-shape. As a consequence, one expects to observe viscous features in transport at intermediate temperatures and densities. The hydrodynamic theory describes the system in general by means of three coupled macroscopic currents (charge, energy, and quasiparticle imbalance current) [11, 12, 80]. As discussed in Section 2.1.7, two different types of hydrodynamic regimes have to be distinguished: the Fermi liquid regime (FL) when $E_F \gg k_B T$, and the Dirac fluid regime (DF) regime when $E_F \ll k_B T$.

In the degenerate limit of the Fermi liquid away from the CNP only a single carrier type is present and all three hydrodynamic currents are equivalent. The system can then be described solely by the macroscopic charge current. Due to the symmetry of graphene's electronic structure, Umklapp processes are inhibited which ensures that electron-electron (or hole-hole) scattering is momentum conserving (and so current conserving). This promotes collective behavior [211]: Any obstacle, residual impurity, or boundary is circumnavigated by this behavior allowing superballistic transport through constrictions at intermediate temperature [205, 212]. Moreover, an injected current will induce vortices (as in a whirlpool) in the highly viscous electronic system [210, 211]. The qualitative agreement of the measured nonlocal resistance map (Figure 2.8c) and the theoretically predicted hydrodynamic window (Figure 2.9a, calculated for an arbitrary impurity density) is striking.

At the CNP electrons and holes coexist and form a strongly interacting quasi-relativistic electron-hole plasma, the so-called Dirac fluid (in theory for ideally clean graphene) [11, 12, 79, 213]. We emphasize that such regime can only be clearly established in real devices when thermally activated charge carriers have flattened the disorder induced inhomogeneous landscape and the system can be considered as homogeneous at charge neutrality (see also Figure 2.9a). As shown in Table

2.1 of Section 2.2.3, the disorder induced potential fluctuations lead to substantially smaller Fermi energy fluctuations in bilayer than in monolayer graphene (due to the different DOS), and so the critical temperature is much lower in the former. Clear experimental evidence of electron-hole collision limited (diffusive) transport was thus first observed in charge neutral bilayer graphene, above the disorder temperature and below the onset of phonon scattering [81]. Equivalent claims for monolayer graphene are still under debate [79, 80]. The system at charge neutrality can be described by two non-equivalent currents in the hydrodynamic picture: a charge current linked to velocity and an energy current being proportional to the net momentum. Electron-hole collisions conserve the latter but cause velocity (and, therefore, current) relaxation limiting transport [11, 79, 81]. The conservation of the energy current, however, has profound consequences for (electronic) thermal transport. When an electric field is applied to the charge neutral hydrodynamic system, electrons and holes drift in opposite direction and will collide which leads to finite resistivity. In contrast, a thermal gradient moves both carriers in the same direction, implying no friction and thus infinite (electronic) thermal conductivity [11, 79]. Hence, a violation of the Wiedemann-Franz law is expected. It was recently reported for high quality monolayer graphene at the CNP in a narrow temperature range (see Figure 2.8b and [79]). We note that the finite resistivity (and its density and temperature dependence) in that experiment can be explained also by the universal disorder limited model which questions the interpretation of e-h scattering limited transport in monolayer graphene [80]. The thermal transport part of the experiment can, however, only be understood in the Dirac fluid regime. The contribution of e-h collisions in thermal transport in graphene with extremely low extrinsic disorder is further shown in experiments which observe an enhanced thermoelectric power larger than the Mott relation at high temperature [188].

Magneto-transport in the hydrodynamic regime has not yet been addressed by experiments but several recent theoretical studies identify possible signatures, e.g. the Hall viscosity [209, 214, 215]. As depicted in Figure 2.8a, the magneto-resistance and the Hall resistivity at finite density in the hydrodynamic Fermi liquid regime can be distinguished from ballistic features [209]: The MR curve does not exhibit a peak feature but decreases monotonically, and the sign of $\Delta\rho_{xy}$ constitutes clear evidence for hydrodynamic effects: At small fields the Hall slope is smaller (larger) than at large fields in the hydrodynamic (ballistic) regime. At the CNP, the Hall voltage is zero and a quadratic positive magneto-resistance arises from hydrodynamic theory [11]. However, when the sample width is comparable with the e-h recombination length, non-saturating linear MR is expected in classically strong magnetic fields [11, 216], and has been observed in bilayer graphene [217].

2.5. Electronic transport in a nutshell and comparison with competing materials

As we have shown in this chapter, electronic transport in graphene is very versatile. The absence of a band gap, its electron-hole symmetric linear dispersion, and the chiral nature of its charge carriers make it a unique material. As a consequence of chirality, (direct) backscattering by long range disorder is (fully) suppressed when the disorder mechanism conserves the pseudospin. Moreover, electron-phonon coupling is weak and optical phonon branches appear at high energy. These two ingredients allow to reach record room temperature mobility values while the average density can be tuned easily at the turn of a voltage knob to very low (ambipolar) values (see Figure 2.9b). Therefore, low disorder graphene is a promising candidate for many applications, e.g. in magnetic field sensing as discussed in this work.

However, this two-dimensional all-surface material is never perfect in experimental reality. While its intrinsic impurity density (e.g. lattice defects) is extremely low, its transport properties are governed by external influences, such as charged impurities in its environment and random strain fluctuations induced by the underlying substrate. This extrinsic long range disorder not only dominates the scattering over a wide parameter range but also imposes a spatially inhomogeneous landscape of e-h puddles with quasi-transparent boundaries when the average density is tuned to zero. As a consequence, the conductivity at the CNP remains finite and has a non-universal (sample-dependent) value in the low temperature limit which can be understood in the framework of semiclassical (self-consistent) transport theory that accounts for the spatial inhomogeneity. Intrinsic Dirac point physics beyond the diffusive picture is then obscured by this extrinsic disorder. Moreover, electrons and holes always coexist at the CNP, either induced by the disorder or thermally activated. In the latter case the system turns homogeneous (above the disorder equivalent temperature T_c). The transport is characterized by a gate-tunable crossover between the two carrier type and the single carrier type regime which corresponds to a crossover from a metallic-like (at high density) to an insulating-like temperature dependence of the conductivity. The coexistence of electrons and holes gives rise to strong classical magneto-resistance which might offer an alternative to measure magnetic fields with graphene (see Section 4.5).

Figure 2.9a summarizes qualitatively the status quo theoretical picture of electronic transport in graphene. The white line demarcates equal contributions of charged

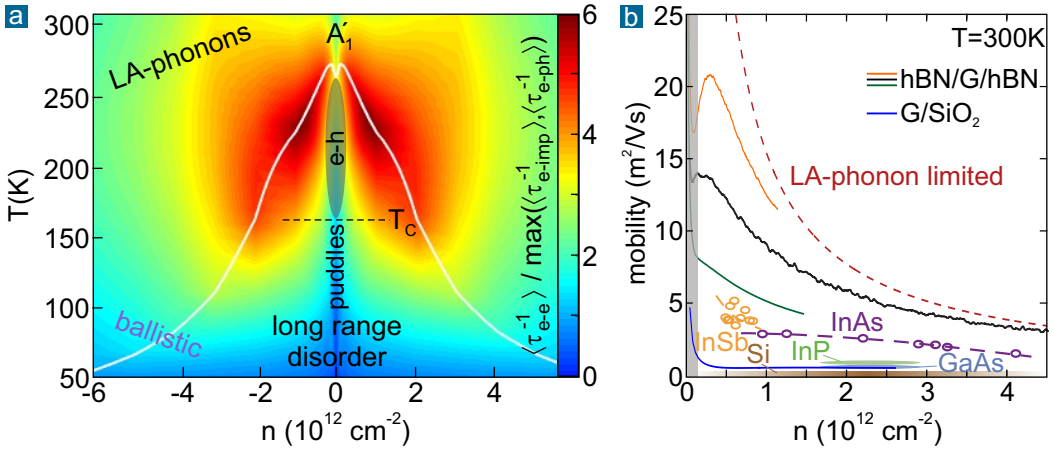

Figure 2.9. Electronic transport regimes in graphene at a glance. Its record room temperature mobility is compared here to common semiconducting materials which are typically limited by strong optical phonon scattering. a: The underlying color plot, taken from theoretical work of Ref. 80, compares the thermally averaged carrier-carrier scattering rate (e-e, h-h, and e-h contributions) to that of charged impurities (as exemplary long range disorder) and all relevant intrinsic phonons (accounting for acoustic and A_1' modes). The white line indicates equal contribution of impurity and phonon scattering. For low enough charged impurity density (or in general long range disorder strength), a collision dominated hydrodynamic window can emerge which is "lung-shaped" for monolayer graphene. Electrons and holes always coexist at low density (close to the CNP). They are spatially separated into puddles by long range disorder fluctuations at low temperature. Thermal activation eventually leads to a homogeneous charge neutral system at high temperature. We have exemplarily drawn this crossover into the graph (T_c in panel a). Only above this temperature, e-h scattering *can* (if at all) become the dominant scattering mechanism in the then homogeneous system before phonon scattering sets in (Dirac fluid regime, indicated by gray ellipsoid). Ballistic transport is typically found at low temperature and high density but its exact boundary is determined by the system dimensions. b: Field effect room temperature mobility of hBN encapsulated graphene and graphene on SiO_2, compared to competing semiconducting materials and to the theoretical longitudinal acoustic phonon scattering limit (based on eq. 2.55 with $D_A = 20\,\text{eV}$). The mobility definition (Eq. 2.38) becomes invalid close to the CNP (gray shaded area at low density). Data points of the black line and of all other materials are taken from Ref. 10, while blue (SiO_2), dark green, and orange traces are our devices presented in Chapter 4. It is obvious that only hBN encapsulation unleashes graphene's full potential.

impurity and phonon scattering. These contributions[56] have been calculated here for an arbitrary but low extrinsic impurity density [80]. In such high quality devices,

56 The theoretical model, presented here, assumes charged impurity scattering as dominant long range scattering mechanism [80]. Qualitatively, all conclusions made here also hold for random strain fluctuations as main long range disorder mechanism. In the following chapters, we

the semi-classical picture of diffusive transport reaches its limits and many-body interactions[57] have to be considered. Carrier-carrier interaction plays an important role at intermediate temperatures and densities. In the single carrier type regime, these interactions result in a collective, fluid-like behavior of the system which is described by hydrodynamic quantities. In contrast, the purely ballistic transport regime is reached (depending on sample dimensions) only at low temperatures and high densities (when carrier-carrier interactions can be neglected).

In the vicinity of the CNP, transport remains consistently diffusive over the whole temperature range. At low temperature, the system at the CNP is spatially inhomogeneous and scattering caused by extrinsic disorder dominates. When the temperature is increased above T_c, thermally activated carriers smooth out the inhomogeneous landscape imposed by the extrinsic disorder. In the conventional picture, scattering in this homogeneous charge neutral system is then governed either by extrinsic disorder or, ultimately, by phonon scattering at high temperature. However, for a homogeneous charge neutral system, the hydrodynamic picture also suggests electron-hole collisions as main scattering mechanism in an intermediate temperature range above T_c and below the onset of phonon scattering. Since both e-h collisions and disorder scattering lead to diffusive transport, the relevance of e-h collisions is still under debate [79, 80].

The map in panel a of Figure 2.9 displaying the different transport regimes is not universal. Every real graphene device is characterized by its unique long range disorder level and so the white line will shift upwards and sidewards for increasing disorder strength. T_c will become larger and will typically exceed room temperature in the majority of devices. Then transport in graphene is completely described by semiclassical theory and, from an experimentalists point of view, a phenomenological model assuming a residual carrier density at the CNP can be applied to analyze the measured magneto-transport data (for classically weak magnetic fields). For commercial sensing applications, the working temperature will be around room temperature which allows us to neglect viscous or ballistic transport effects. However, only extraordinarily clean devices with low extrinsic disorder can be considered as homogeneous charge neutral systems at zero average density around room temperature. For most scalable manufacturing techniques, the extrinsic disorder level will be so large that the system will be spatially inhomogeneous in the temperature range of commercial applications.

will distinguish between contributions of extrinsic disorder scattering (in general) and phonon scattering. High quality devices are then defined by particularly low extrinsic disorder.

57 Many body interactions also cause a renormalization of the Fermi velocity and so change the band dispersion in graphene. This substantially influences thermal charge carrier generation.

Chapter 3

Pushing graphene devices towards intrinsic limits

This chapter gives insight into the production, handling, and device processing of 2D materials, with particular focus on graphene and hexagonal boron nitride as its ideal encapsulation. In the former chapter we have highlighted that graphene's electronic properties are mainly governed by extrinsic sources of disorder. Both direct probe (Section 2.2.3) and transport experiments (Section 2.3.2) have revealed that the main culprits are charged impurities both in close and remote distance to graphene and the substrate's surface roughness leading to random strain fluctuations. The experimental endeavor to push graphene towards its intrinsic limits has been one of the main tasks of this work. The development of our transfer setup and stacking procedure was thereby sparked by key achievements and proposals within the 2D materials research community, especially the use of hBN as a substrate and the van der Waals heterostructure assembly technique [10, 161]. Furthermore, we have developed a reliable, post-processing technique for improving the quality of such heterostructures and increasing the yield of excellent devices. The main results of this chapter - the description of our proposed routine, our hypothesis of the underlying mechanism, and the statistical evidence for its beneficial effect - have been also published in Ref. 183.

The chapter is split into two main sections. The first section gives an overview of the different device processing routines which can be distinguished according to their strategic approach to minimize extrinsic impurities. These impurities are either reduced subsequent to the device processing or avoided for the most part through encapsulation of the graphene layer in the first step. In the first approach, graphene's surface is initially (allowed to be) exposed to all processing steps (resists, solvents, wet lift-off of metals, and even etchants) and the extrinsic impurity density is then lowered by cleaning treatments. This approach is versatile and can be adapted also

to large scale production methods such as chemical vapor deposition (CVD) growth of graphene. This technique allows to transfer graphene (or another 2D material) onto almost any substrate, e.g. atomically flat hexagonal boron nitride. The second approach is the van der Waals heterostructure assembly method [10]. Here, graphene (or another 2D material) is encapsulated between other thicker layered crystals, typically hBN. The latter serves as protective environment, before any further device processing steps are conducted. As a big bonus, self-cleansing effects occur between certain types of layered crystals with high affinity to each other which ensures atomically flat and ultra clean interfaces. The resulting improvement in the electronic properties is unsurpassed by any other production technique. Hence, van der Waals stacking has recently become the workhorse in basic research of 2D materials. The choice of substrate materials, however, is restricted to only a few layered crystals for which self-cleansing takes places. As of today, the stacking method cannot be adapted to large scale production, with the lateral dimensions of flat and stepless hBN crystals being the main bottleneck. Finally, we present a routine that allows obtaining highest quality heterostructures more reliably by combining thermal and (contact AFM based) mechanical treatments.

In the second section we report on our device statistics illustrating the difference between SiO_2 and hBN as a substrate as well as the gain in quality of hBN-encapsulated devices after applying the here proposed technique. We also compare the above mentioned different device production methods. Our main figures of merit for electronic transport are the residual carrier density at the CNP and the charge carrier mobility which both are determined by the extrinsic disorder level. We also consider initial doping level and hysteresis effects upon gating when comparing devices on SiO_2 and hBN. These findings are then put in relation to the investigation of a large ensemble of graphene samples by Raman spectroscopy. The latter not only allows to identify unambiguously the layer thickness but is also sensitive to strain and doping of the graphene lattice. Hence, Raman spectroscopy constitutes an important non-invasive tool to estimate and monitor the electronic properties before and during device processing. However, we also demonstrate that Raman spectroscopy approaches its limits to distinguish between different grades of high quality, hBN-encapsulated devices. Then, one relies on electrical transport measurements which are more sensitive to the device quality in terms of very low extrinsic disorder level. We provide criteria to identify promising hBN-encapsulated graphene devices already at room temperature and show that their room temperature properties are then mainly governed by intrinsic limits, i.e. phonon scattering at high densities and thermal activation of charge carriers at the CNP.

3.1. Device fabrication - eliminating extrinsic disorder

On the scope of this section

In the following section we address solely *mechanical exfoliation* of bulk layered crystals by adhesive tape to obtain (atomically) thin sheets of 2D materials, such as graphene and hexagonal boron nitride in this work. Up to now, this manufacturing method produces graphene with the lowest intrinsic disorder level, e.g. with respect to vacancies and lattice impurities. Thus, it serves as benchmark for other, scalable production techniques. Moreover, we restrict us to *substrate supported graphene devices*, since freestanding (or suspended) graphene exhibits various technical constraints. The device production process is very delicate and a high failure rate is unavoidable. The device dimensions are restricted to a few micrometer (or even below) and typically only two-terminal or unconventionally cross-shaped geometries can be used [196, 218, 219]. Especially, the current annealing technique which is used to burn off very efficiently extrinsic impurities from the device works only properly for such geometries [218]. Although such current annealed suspended devices feature excellent electronic properties at low temperature holding the record of $< 10^9\,\mathrm{cm}^{-2}$ charge inhomogeneity at CNP [126], their room temperature mobility is limited to $10,000\,\mathrm{cm}^2/\mathrm{Vs}$ due to the presence of flexural phonons [130]. In conclusion, suspended graphene is very unlikely an aspirant for mass scale production of complex devices, e.g. Hall bars, and for room temperature applications. In the framework of basic research, however, it allows to study graphene's properties without interaction with a substrate, e.g. Moire effects on hBN.

A short comment on other ways to produce (large-scale) graphene

Manifold approaches towards large scale fabrication have been proposed which are reviewed in Refs. 220, 221. The bottom-up chemical synthesis of graphene nanoribbons [222, 223] and the liquid phase or chemical exfoliation method [224] are only of interest for applications that do not rely on graphene's record transport properties. Wafer scale graphene of reasonable electronic quality can be epitaxially grown on SiC by selective thermal sublimation of Si atoms from the surface [225, 226]. This approach is, however, restricted to SiC having a terrace-like surface from which graphene cannot be transferred to an arbitrary substrate. Although an effective decoupling of the graphene layer from SiC by hydrogen intercalation solves the issue of strong initial hole doping [227], the room temperature mobility still remains around $5000\,\mathrm{cm}^2/\mathrm{Vs}$ [228, 229]. The latter does not favor the use of epitaxial graphene in

electronic applications. However, a peculiarity of epitaxial graphene on SiC, resulting from strongly localized bulk states, makes the system an interesting candidate as quantum Hall resistance standard requiring lower magnetic fields than GaAs [228].

In contrast, there has been substantial progress in the development of chemical vapor deposition (CVD) growth and transfer routines, not only for graphene but also for other 2D materials such as hBN and transition metal chalcogenides [230, 231]. Growth of single crystalline graphene with lateral dimensions of at least some millimeters and even up to wafer scale has been demonstrated on various catalytic substrates, e.g. hydrogen terminated germanium [232] and copper (alloys) [231, 233–235]. The main advantage is that the CVD grown material, e.g. graphene, can be transferred to arbitrary substrates. It has been found that the electronic properties of CVD grown graphene correlates with the choice of the target substrate and with the cleanness of the transfer process, in analogy to exfoliated graphene. This allows to adapt improvements in handling and processing of the latter, and vice versa. Hence, we will point out these synergies in the following section. Today's electronic properties of CVD grown single crystalline graphene have reached the level of exfoliated flakes when the grown layer is encapsulated between exfoliated hBN flakes. Accordingly, room temperature mobilities up to $100,000\,\mathrm{cm}^2/\mathrm{Vs}$ have been achieved [236]. This implies that the bottleneck for industrial scaling is not anymore growing large single-crystalline graphene but providing ideal substrates (and encapsulation) with wafer-scale dimensions. Hence, the development of hBN growth techniques has gathered momentum in the recent years, but the results hitherto have not been disruptive and we still rely on exfoliated hBN flakes for proof of principle studies.

3.1.1. Fabrication of (bare) graphene devices with uncovered surface

Layered crystals such as graphite are only weakly held together by van der Waals forces. They can be thinned down to a single layer through micromechanical cleavage using adhesive tape[1] which is referred to as mechanical exfoliation technique [1, 14]. The tape will initially pick up a rather thick and dense film from the bulk crystal. The coverage and thickness can be reduced by either folding the tape a few times or applying several fresh pieces of tape to the first tape (typically 3-5 iterations). This procedure ensures that the majority of the flakes on the tape has a freshly cleaved surface. Then, the tape is pressed onto a suitable substrate, e.g. a piece of silicon

1 Materials and sources: Natural graphite flakes (NGS Naturgraphit GmbH), hBN flakes (HQ-Graphene and from collaborators at NIMS Japan), 3M Scotch Magic tape for graphite, and Nitto BT-150E-CM processing tape for hBN.

wafer with an oxide layer. Many flakes of different layer thickness remain stuck on the substrate's surface after the tape has been slowly pulled off again. Choosing the right silicon oxide layer thickness is thereby the key enabler to make even a single atomic layer visible under an optical microscope due to contrast enhancement by an interference effect [83]. A trained eye is then capable to distinguish between mono-, bi-, and multilayer graphene (see Figure 3.1 top-left panel). We either use wafers with 90 nm oxide thickness (from UniversityWafer Inc., to obtain flakes for pick-up) or wafers with a dry thermal oxide of 290 nm thickness (from IMS Chips, used for all devices in this work). The latter silicon wafers are strongly p-doped with boron and exhibit a backside metallization which allows us to use them as back gate for our devices. A marker system is produced by electron beam lithography prior to exfoliation such that the randomly distributed flakes found by manual search can be reliably mapped and relocated in the subsequent characterization and processing steps.

The graphene layer thickness can be unambiguously determined by Raman spectroscopy (see Figure 3.5). Atomic force microscopy (AFM) is used to measure the thickness of hBN flakes. For reasons we will discuss below, the hBN thickness should be between 15 and 40 nm. Under an optical microscope this corresponds to bluish (slightly transparent) and matt green colors, respectively. Since the flakes are freshly cleaved when the tape is pulled off from the substrate, their surface is free from any tape residues. In contrast, the surrounding areas of the SiO_2 surface have been in direct contact with the adhesive tape and often exhibit gluey residues. The bare graphene devices with an unprotected surface are then manufactured by electron beam lithography which involves covering the graphene flake at least twice with PMMA resists, thermal treatments during resist curing (at 160°C), and solvent based removal of the PMMA layers (see Figure 3.1 top-right panel). In the first step, a PMMA layer serves as mask to define the geometry of the device through oxygen plasma etching. After the successful etching of the graphene layer, the PMMA layer is removed in warm acetone (ACE) followed by a dip into isopropyl alcohol (IPA) after which it is dried with a nitrogen gun. In the second step, a double-layer of PMMA is spincoated and electrical contacts partially overlapping the graphene are designed. Typically we evaporate chromium/gold contacts where the thin chromium film (about 5 nm) acts as adhesion layer. The metal lift-off process is performed in warm ACE, followed by the same routine as above. It is obvious that these processing steps will have a detrimental impact on non-encapsulated graphene. Thermal treatments lead to changes in the strain distribution, while PMMA residuals and metallic particles adhere to graphene constituting a major source of extrinsic impurities and local strain. Several post-processing cleaning routines have been proposed to address these issues which we will briefly discuss in Section 3.1.2.

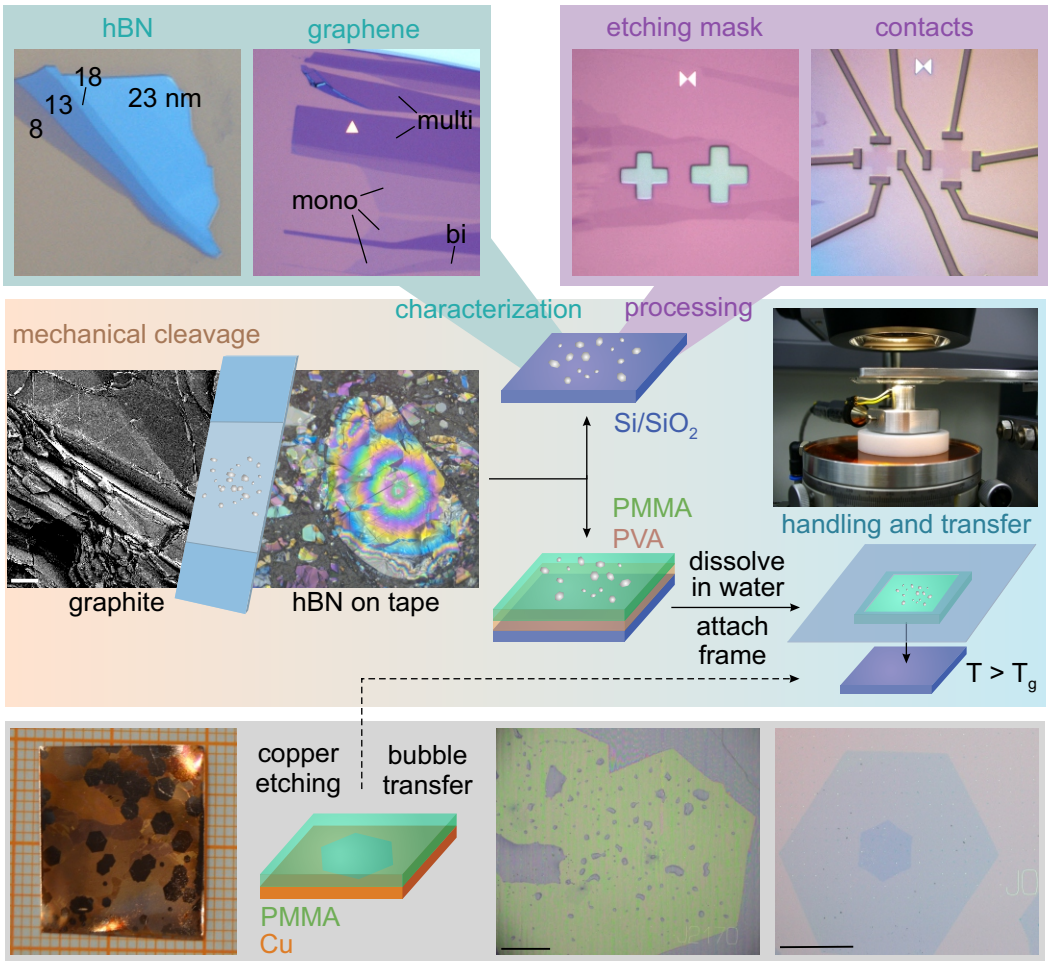

Figure 3.1. Device fabrication of non-encapsulated, bare graphene (and other 2D materials). Middle panel: 2D materials are cleaved from their bulk crystal (SEM picture of graphite, scale bar 100μm) using a tape (optical picture of thinned hBN crystals on tape) and finally transferred to the target substrate, e.g. SiO_2/Si. On the latter, layer thickness can be characterized via optical microscopy, Raman spectroscopy, and AFM height measurement (top left panel). Devices are fabricated by electron beam lithography which exposes the 2D material surface to resists and solvents (top right panel, marker size 5x5μm). Exfoliation onto polymer double layers allows us to handle and transfer 2D crystals onto arbitrary substrates or on top of each other by a transfer setup (picture). Polymer film assisted transfer is also used in case of our CVD grown graphene on copper, as shown in the bottom panel (scale bars are 500μm). See main text for details.

The mechanical exfoliation technique has two major drawbacks. First, its yield is low and the typical area of monolayer flakes is restricted to a few hundreds of

μm^2. Second, the flakes are randomly distributed and monolayer graphene is not visible on arbitrary substrates. Solutions to these problems are presented in the next paragraphs.

Effect of SiO_2 substrate treatments

Increasing the yield is directly linked to the complex question about the microscopic nature of the forces between graphene and the functional groups which terminate the surface of the amorphous SiO_2 substrate. Additionally, both surfaces are exposed to airborne contaminants, such as hydrocarbons and atomically thin water films. Until today, it is still under debate whether graphene is hydrophilic or hydrophobic [237, 238]. There is experimental evidence by time-dependent water contact angle measurements that freshly cleaved surfaces are at first less hydrophobic becoming then more hydrophobic with time due to airborne hydrocarbons [237]. Further, the functional groups of the SiO_2 surface can be modified by different pre-treatment steps [94, 239], e.g. oxygen plasma treatment, hydrofluoric acid etching, and attachment of self-assembled monolayers (SAM) of hydrophobic molecules [139, 240–242].

Based on our experience, it is difficult to relate a certain SiO_2 substrate treatment to the final device performance (see Section 3.2). We have found that electronic properties within one batch (subject to the same treatments) can vary to the same extent than between different treatments. However, three empirical conclusions can be made: First, the yield of exfoliated flakes possessing a large area beyond thousand μm^2 is substantially increased by fresh oxygen plasma treatment of the SiO_2 surface prior to exfoliation, since this removes efficiently hydrocarbons (and residues from the marker system production). The water contact angle is zero right after this treatment which demonstrates that the surface is extremely hydrophilic (due to silanol groups). Within one day at ambient conditions the contact angle increases back to the untreated value, in agreement with observations of other groups [94]. Second, we achieved better exfoliation results when the substrates were pre-heated to 120°C on a hot plate and the tape was applied during cool-down of the substrate (as also reported independently in Ref. 243). The latter routine allowed to omit the fresh plasma treatment before exfoliation while still maintaining a good yield. The two pre-treatment routines (fresh plasma vs 120°C pre-heat) lead to a statistically notable difference when studying the Raman spectra of a large ensemble of graphene flakes after exfoliation (Section 3.2.1). However, after device processing we cannot differentiate anymore between these two routines which indicates that the processing steps determine the final properties. Third, several studies report that

hydrophobic SAM films on SiO_2 improve the electronic properties, e.g. HMDS [240], OTS [139], or APTES [242]. Both graphene's initial doping level and the hysteresis in gating are found to be reduced compared to devices on bare SiO_2. The authors also claim a slight improvement in the mobility by a factor of two which however still does not exceed $10,000\,cm^2/Vs$ - a rather poor value with respect to much higher mobilities which are reliably achieved on hBN. Since the hydrophobic treatment, in our experience, lowers significantly the exfoliation yield and complicates device production, e.g. spincoating resists, we have not pursued this approach.

In summary, higher yield in mechanical exfoliation is achieved by a stronger interacting, hydrophilic SiO_2 surface which is free of hydrocarbons right after plasma treatment. As a trade-off, graphene's electronic properties are more influenced by the substrate to which it adheres firmly [244]. The increased attraction of the plasma treated SiO_2 surface can be explained by negatively charged silanol groups which are not saturated anymore by hydrocarbons [94]. Such strong interactions can also attract other adsorbates due to the increased chemical reactivity [95, 139].

Polymer assisted transfer technique

Supporting graphene or other 2D materials by polymer films has enabled us to handle these fragile layers safely and transfer them reliably with high spatial precision onto arbitrary substrates. For exfoliated flakes, we start by preparing a stack of two different polymers onto a bare SiO_2/Si wafer piece, choosing the layer thicknesses such that the contrast enhancing interference effect persists. Following the proposal of Ref. 161, the bottom layer is water soluble (poly-vinyl alcohol), while the top layer later supporting the flakes is not. For the top layer one can choose from a multitude of polymers with respect to glass transition temperature, exfoliation yield, substrate-solvent compatibility, and easiness to remove them residue-free. We use PMMA as top polymer since the exfoliation yield is high and even larger monolayer flakes than on plasma treated SiO_2 can be obtained when the wafer with PVA/PMMA double layer is pre-heated to 120°C prior to exfoliation. PMMA and solvents are anyway used during device processing which limits the number of different substances in contact with graphene. After searching and identifying suitable flakes by Raman spectroscopy, we attach a plastic frame to the PMMA film with double-sided tape (see Figure 3.1). The PVA layer is then removed in DI water and the PMMA film delaminates from the SiO_2/Si wafer piece. The frame stabilizes the PMMA layer and allows to remove it without fracture or wrinkling from the water beaker. After drying, the frame with the PMMA film is mounted upside-down to a holder which is connected to a z-axis micrometer positioning stage. This allows to

slowly approach the PMMA film to the target substrate, for example a hBN flake exfoliated onto another SiO_2/Si wafer piece. The target substrate is fixed to a metal block with integrated heater which can be positioned in xy-direction. By means of a long working distance microscope objective the flake on the PMMA film can now be transferred precisely onto the target flake or any desired substrate position. A photograph of the whole transfer setup is shown in Figure 3.1. To achieve a stable and smooth contact between the PMMA film and the target substrate, we increase the temperature slightly above the glass transition temperature of the polymer (for our PMMA $T_g \approx 130°C$). When the glass transition temperature is reached, a rapidly propagating change in color (from dark gray to greenish) indicates that the PMMA film (and so the flakes) are attached properly. After cool-down, the PMMA film is cut off from the frame and the PMMA is removed with the standard ACE/IPA wash.

The method of supporting graphene with a polymer film is also widely employed for transferring CVD grown graphene from the catalytic metal foil to a target substrate (see Figure 3.1 bottom panel and the review in Ref. 245). For instance, we have grown millimeter sized single-crystalline graphene hexagons on poly-crystalline copper foils[2]. Subsequently, a PMMA layer is spincoated on the copper foil in order to protect the graphene film when the copper is etched away in a wet chemistry process [245, 246]. Alternatively, the PMMA film together with the grown graphene is delaminated in an electrochemical process (referred to as bubble transfer [249, 250]). Typically, the floating PMMA film, supporting the graphene, is then directly fished out of the beaker with the target substrate (wet transfer). We have modified this procedure by attaching a frame to the PMMA film prior to the copper etching (or delamination) step. Hence, we are able to handle the PMMA film in a similar manner as described above for exfoliated flakes. In a dry process, we use the aforementioned setup to transfer CVD graphene to our SiO_2/Si substrates. In the microscopic images shown in the bottom panel of Figure 3.1, one can see nicely that the PMMA part with (hexagonal) graphene already tends to stick at lower temperature to the plasma pre-treated wafer, in contrast to the bare PMMA film. This is another argument favoring the hypothesis that graphene is attracted by a hydrophilic SiO_2 surface (see previous paragraph). When the temperature is increased, the whole PMMA film finally attaches, and after its removal we obtain millimeter sized single-crystalline monolayer (and bilayer growing with slower rate from the same nucleation center) graphene on the target substrate.

2 Our work on CVD growth of graphene will not be covered by this thesis. For a general introduction to CVD growth of graphene (and other 2D materials), we recommend Refs. 230, 232–235, 246–248.

As we have demonstrated, the development of handling and transferring routines is of general and fundamental importance for both exfoliation and CVD growth of 2D materials. Optimizations, e.g with respect to the used polymers and to engineering tricks in the transfer process or setup, can be adapted easily among both production methods which creates valuable synergies. Graphene has been spearheading the development of these techniques which are now available to many other 2D materials speeding up their research progress.

3.1.2. Cleaning graphene's surface from extrinsic impurities

In this paragraph, we address the question how one can efficiently clean graphene devices with exposed surface from extrinsic impurities - a truly Sisyphean task. To give an answer, we have to classify potential extrinsic impurities and their origin. The first type we classify as "bulk" (mesoscopic) stemming from device processing, e.g. PMMA or metallic residues (including copper residues in case of CVD graphene). They can be seen sometimes even under an optical microscope or at least have several nanometer height such that they can be observed in non-contact AFM having no atomic resolution (see Figure 3.2c). Moreover, they are found to be ubiquitous in low magnification TEM images of polymer based transferred graphene making it difficult to find a spot for unimpeded imaging [251, 252]. The second type are airborne adsorbates, such as hydrocarbons [237] or water and oxygen molecules [95, 253], which are expected to cover the whole graphene surface (and likely also the interface between graphene and its substrate). They are blamed to cause initial doping and contribute to the hysteresis in gating [95, 240]. Their density and adhesion strength is related to the reactivity of graphene's surface which ultimately is determined by the substrate material and the degree of graphene's conformation to the latter [95, 136, 139].

Many ideas to clean graphene have been raised and their effect has been studied with different ambitions [9]. For TEM, the target is to obtain locally clean areas on free-standing graphene for imaging purposes. In contrast, electrical transport properties are determined by the average global density of both types of extrinsic impurities, and such devices are typically supported by a substrate. The most frequently proposed cleaning methods (for PMMA) are listed here:

- stronger solvents such as NEP, NMP, or acetic acid [254]
- current annealing (ohmic heating by a large current) [218, 255, 256]

- hydrogen plasma treatments [257]
- mechanical cleaning with an AFM tip [258–260]
- thermal annealing in a reducing atmosphere [161, 251, 252, 261, 262]

Several of these studies do not measure electrical transport, but monitor the removal of surface contamination by polymer residuals via TEM imaging, x-ray photoelectron spectroscopy (XPS), electron energy loss spectroscopy (EELS), or AFM scans. Graphene is considered as clean when a freestanding flake in TEM shows several hundreds of nanometer sized areas without polymeric residues [251, 252, 261]. XPS spectra reveal the vanishing of covalent carbon bonds assigned to polymeric residues and an increasing clean graphene related carbon sp^2 peak after thermal and plasma treatments [251, 257, 262]. AFM measurements on supported graphene show smoother areas with lower density of "bulk" residues subsequent to cleaning treatments [254, 255, 257–260].

However, these findings do not allow us to draw conclusions regarding the final transport properties of non-encapsulated graphene devices on substrates which have been exposed to ambient conditions prior to measurements. In this case, airborne adsorbates (the second type of extrinsic impurities) can act as dopants by charge transfer and as scattering sites. These contaminants cannot be detected by AFM measurements[3] but they will have a drastic (detrimental) impact on transport and Raman spectra. Moreover, these physical adsorbates do not form covalent bonds to the graphene lattice such that XPS spectra indicate "clean" graphene with pure sp^2 bonds despite their presence.

We have tested all except the plasma cleaning method on several non-encapsulated devices supported mainly by SiO_2 (and some by hBN). Their cleaning effect regarding polymeric residues are monitored by AFM measurements, while statistical studies based on Raman spectroscopy and transport measurements give insight into the effect of molecular adsorbates (see Section 3.2). We find these methods, if at all, to clean graphene only from "bulk" impurities, while especially thermal annealing even impairs electronic properties. After device processing and cleaning in warm acetone, graphene on SiO_2 exhibits many spike-like features of residues seen in AFM scans (Figure 3.2c). In our experience, the use of stronger solvents neither substantially reduces these "bulk" residues nor has a statistically significant effect on the electronic properties (in comparison to the standard acetone PMMA removal). We

3 And they are not seen in TEM since freestanding flakes are typically annealed in UHV such that these volatile adsorbates are removed before imaging.

find the current annealing method to not work properly for larger areas and for complex device designs with more than two-terminal contacts (not shown). In addition, this technique suffers from a high device failure rate and is not scalable.

The lack of scalability is also a major drawback of the mechanical cleaning method which, however, is found to reliably brush away all residues within the scanning range (Figure 3.2c). A non-contact mode tip with large tip diameter (about 20 nm) is used to scan the target area in soft contact mode (force is set to 0 to -20 nN). The tip piles up the dragged residues at the return points of the scan. This opens a clean window of non-damaged graphene on SiO_2 possessing a surface roughness of only 155 pm, comparable to bare SiO_2. Raman spectra taken before and after AFM cleaning reveal a slight reduction in the doping level, in agreement with transport measurements reported in the literature [258–260]. The initial doping level shifts from substantial hole doping to almost zero, whereas the improvement of the mobility is at most a factor of two and values remain below $10,000\,\mathrm{cm}^2/\mathrm{Vs}$. This indicates that polymeric residues are not the main cause for scattering on SiO_2.

Thermal annealing effects on non-encapsulated graphene

Thermal treatment of graphene devices has become a standard and scalable procedure for cleaning [94, 161, 262]. We perform two different thermal treatments on our non-encapsulated, substrate supported devices. The first treatment is optional and conducted in a rapid thermal annealing (RTA) oven in a continuous flow of forming gas ($10\%\,H_2$ in Ar) at fixed pressure of 150 mbar. The temperature is increased gradually to $300 - 350°C$ and held for at least one hour before starting a slow cooldown ramp. This temperature range is at least required such that PMMA residues are efficiently removed by breaking the polymerization backbone bond while still keeping the graphene lattice intact [251, 262]. The samples are then re-exposed to ambient conditions (e.g. moist air) before Raman or transport measurements are conducted. While the Raman spectra are taken in ambient conditions, electrical transport is typically probed in high vacuum ($p \approx 10^{-4} - 10^{-5}\,\mathrm{mbar}$). The sample holder for electrical measurements can be heated to 140°C which allows us to bake the devices in-situ and in vacuum (our second thermal treatment step).

The high temperature annealing in forming gas notably burns off most AFM detectable polymeric residues ("bulk" impurities), leaving behind only few leftovers but causing graphene on hBN to exhibit wrinkles and bubbles (similar to those reported in Section 3.1.4). However, we find the electronic properties of graphene on SiO_2 to be drastically affected by this treatment. As we report in detail in Section 3.2, both

Raman spectroscopy and electrical transport measurements reveal a much stronger hole doping (and hysteresis) than for untreated devices. This doping level persists even after the 140°C vacuum annealing step in the electrical sample holder, in contrast to devices which have not been annealed at 350°C. We also annealed freshly exfoliated flakes (not subjected to device processing steps, e.g. PMMA) and found the same doping behavior by means of Raman spectroscopy. Similar observations have been reported in literature performing vacuum instead of forming gas annealing at the same temperature [263]. Hence, a connection of this effect to PMMA or hydrogen can be excluded. We find the strong doping effect to be related to the SiO_2 substrate since we did not observe similar changes in the Raman spectra for hBN supported graphene exposed to the high temperature annealing step (Figure 3.6). This finding is in accordance with elsewhere reported transport measurements on hBN [161] and with studies comparing annealing effects on SiO_2 and suspended graphene by Raman spectroscopy [95]. The authors of the latter study [95] in addition exposed, in an in-situ experiment, the annealed flake selectively to purified gases. They clearly prove that dry molecular oxygen causes the strong hole doping which is reversible by argon flushing. The doping, however, becomes irreversible as soon as oxygen with a trace of water vapor is used (simulating ambient conditions). Following their hypothesis, the water molecules stabilize a nonbonded graphene-oxygen charge-transfer complex. The role of annealing is then that the interface between SiO_2 and graphene is cleaned such that graphene conforms stronger to the substrate which increases its corrugation and chemical reactivity [95, 139].

The strong hole doping on SiO_2 after annealing has meanwhile been confirmed by several other groups [94, 253, 263–265]. However, the exact microscopic mechanism for the hole doping is under debate. All theories involve oxygen species as source for charge transfer. However, oxygen might also be located below graphene since the hole doping can be reversed after immersion of the annealed flake into DI water for several days [253]. It is shown that water molecules intercalate very slowly between graphene and SiO_2, lowering the doping level to the pre-annealed value. Moreover, the increase of the chemical reactivity with stronger corrugation is questioned by another experiment studying the oxidative reactivity of graphene supported by various substrate [136]. Graphene on hBN remains intact and shows only little doping even after being exposed to an oxygen atmosphere at 500°C. In contrast, it is etched and shows strong doping not only on SiO_2 but also on atomically flat mica flakes. The authors thus relate the larger chemical reactivity to the higher charge trap density in oxide substrates. We have obtained similar results when testing thermal cleaning in Ar/O_2 flow at 350°C which etched graphene on SiO_2 but left graphene on hBN undamaged.

Conclusion

The inconvenient truth is that there is no method available to push substrate supported, bare graphene (with uncovered surface) towards its intrinsic limits, e.g. to obtain a room temperature mobility beyond $100,000\,\mathrm{cm}^2/\mathrm{Vs}$. Although some cleaning methods allow us to reduce polymer residues and initial doping levels, an exposed surface will always make the device sensitive to the environment and even a small remaining impurity density on the surface will deteriorate graphene's electrical properties. We do not recommend the thermal cleaning technique performed at 350°C for graphene on SiO_2 since it leads to heavy hole doping after exposure to air. Hence, all our devices on SiO_2 in this thesis were only cleaned by solvents, following the motto "keep it simple" and neglecting the not scalable mechanical cleaning technique (which does not improve the mobility). It is then possible to lower the hole doping level by baking the samples at 140°C in vacuum directly before transport measurements which removes weakly attached volatile adsorbates. With this approach we achieve maximum mobilities on the order of $10,000\,\mathrm{cm}^2/\mathrm{Vs}$ on SiO_2 which tends to be the substrate inherent limit.

Just by placing graphene on hBN, the aforementioned routine already results in average mobilities between $10,000$ and $40,000\,\mathrm{cm}^2/\mathrm{Vs}$ [18, 161], at which annealing at 350°C can be used without the hole doping problem. This demonstrates that an ideal substrate (as defined in the next section) is more beneficial than cleaning treatments performed on a bad substrate. Since graphene is still exposed to fabrication residue contamination in the open surface approach, encapsulation will enhance the mobility even further (Section 3.1.4).

For large area CVD grown graphene we cannot rely on exfoliated substrates such as hBN and (yet) are forced to use non-ideal but wafer-scale substrates such as SiO_2 (with respective mobility limits). However, high temperature annealing, the only widely tested scalable method, is not applicable on these substrates. The next best strategy is to avoid polymer residues and passivate graphene's surface in order to protect it from airborne adsorbates. Following this strategy, several alternative residue-free supporting films [266–269] and even completely support-free techniques [270, 271] have been devised to transfer CVD grown graphene from the catalyst foil. Due to the wafer scale approach, electron beam lithography can be replaced by optical lithography featuring easier removable resists. Alternatively, shadow masks can be used for device processing. Finally, the devices are passivated by resists [272] or oxides [93, 139, 176] to prevent adsorbates from attaching to graphene's surface which fixes the initial doping level.

3.1.3. Ideal substrates as key to high performance graphene devices

The implementation of crystalline hexagonal boron nitride flakes as substrate has been one of the major breakthroughs in graphene research [161]. It changed the community's mindset by revealing that a novel approach employing layered crystals as substrates is the key to unleash graphene's theoretical record properties in real devices. The resulting transport metrics outperform those achieved by former efforts, e.g. optimizing SiO_2 by interface engineering [94, 240] or using alternative amorphous oxide substrates with larger dielectric constant for enhanced charged impurity screening [135, 173, 176, 191]. Two-dimensional, all-surface materials such as graphene have exceptional requirements regarding their environment, in contrast to bulk materials (including other 2DEG embedded in bulk crystals). Scattering in graphene is governed by charged impurities and random strain fluctuations imposed by the substrate roughness (Sections 2.2.3 and 2.3.2). Moreover, charge transport can be impeded by interfacial polar phonon modes at room temperature. This sensitivity to extrinsic sources of disorder is a direct consequence of its two-dimensional nature. A single atomic layer cannot screen properly the effect of charged impurities in its vicinity and it conforms to the substrate surface due to its high out-of-plane flexibility. While random strain fluctuations have a clear origin, the sources of charged impurities are ubiquitous. Trapped charges embedded in the substrate, dangling bonds terminating the substrate surface, and adsorbates adhering directly to graphene or being trapped in between graphene and the substrate are all potential candidates (see also Section 2.2.5 and Figure 2.5).

Hence, to boost graphene's electronic properties, the optimal substrate material is *required* i) to possess a low intrinsic density of charged impurities, ii) to be atomically flat featuring a surface termination without dangling bonds, iii) to show a high affinity to graphene such that self-cleansing effects do occur, and iv) to exhibit highly energetic polar optical (surface-)phonon modes to minimize remote-phonon scattering at elevated temperature, e.g. room temperature. While the first (i) and the last (iv) requirements are self-explanatory, the other two demand further explanation.

Obviously, an atomically flat substrate surface lowers the amplitude of random strain fluctuations and thus the related scattering [18]. More important, graphene and the substrate form a continuously smooth interface without nano-cavities and possible diffusion pathways, e.g. for molecular adsorbates trapped between graphene and SiO_2. The required chemical inertness[4] in terms of absent (polar) dangling

4 Chemical inertness of course also implies that the substrate material should be airstable and non-reactive towards all processing steps. Moreover, it should not establish covalent bonds to

bonds at the surface prevents the otherwise arising charge transfer between graphene and these bonds (compare to the hole doping of graphene on SiO_2 due to polar silanol groups [94, 253]). This requirement further reduces significantly the number of possible sticking sites for airborne contamination (and polymer residues) which makes them footloose and mobile. If the affinity between graphene and the flat substrate exceeds the affinity between them and the contaminants, the latter are pushed away to achieve the largest possible common interface between graphene and the substrate since this condition is energetically favored [138, 273]. This so-called self-cleansing effect leads not only to atomically smooth but also to perfectly clean interfaces. Accordingly, it addresses both major scattering sources, charged impurities and strain fluctuations, at once. The occurrence of this effect has so far been only observed on lipophilic substrate materials [129, 138].

Besides these mandatory requirements on the candidate substrate to obtain good electrical properties, what "soft skills" would be nice-to-have? Since we are interested in electrical transport, the substrate itself should not be conducting. In the ideal case, the material is an insulator featuring a large band gap, a high breakthrough voltage, and a high dielectric constant, such that it can be used as an efficient dielectric for field effect gating. Another constraint might be the visibility of graphene on the substrate, especially when dealing with exfoliated flakes. Moreover, having in mind to exploit the high transport quality of graphene in applications, the material should be scalable towards wafer level. In final device schemes for applications, it might not only serve as substrate but also as encapsulation layer. This naturally raises the question whether it can be grown on graphene (and vice versa), or transferred onto it in a clean process.

What kind of material families can fulfill this long list of requirements? Bulk oxide substrates (e.g. SiO_2, sapphire Al_2O_3, $SrTiO_3$, or hafnia HfO_2) perform poorly. Due to weak bonding in their (amorphous) atomic structure they exhibit a comparatively high intrinsic density of charged impurities. Their surfaces are atomically rough and typically terminated by polar dangling bonds which interact with graphene and contaminants. Hence, no sign of self-cleansing has been observed between graphene and these materials. Finally, the polar optical phonon branches are at low energy and are therefore populated at room temperature. This causes considerable remote phonon contributions to scattering [90, 191]. In contrast, a natural choice are other layered crystals in the family of van der Waals materials [15, 138, 273, 274]. By definition, their individual layers are held together only by weak van der Waals forces instead of covalent or ionic bonds. Accordingly, they possess atomically flat

graphene which would alter tremendously the bandstructure and hence the electronic properties of graphene.

and dangling bond free surfaces. These layered crystals often feature strong in-plane bonding which minimizes substantially the density of charge traps and voids.

A broad library of layered crystals has become available in recent years, mainly by mechanical exfoliation of high quality bulk crystals [15, 273, 274]. Regarding their use as substrate for graphene, however, some of them are favorable while others have turned out to be less promising. Several experiments have studied and compared the effects of layered oxides, transition metal dichalcogenides (TMD), and hexagonal boron nitride (hBN) on the transport properties of graphene [134, 136–138, 138, 275]. Atomically flat oxides, e.g. bismuth strontium calcium copper oxide (BSCCO), vanadium pentoxide (V_2O_5), or muscovite mica, contain a higher density of charged impurities [136] and, more important, they do not show a high affinity to graphene due to their hydrophilic nature. No self-cleansing effect occurs and contaminants remain distributed all along the interface which can be monitored by AFM mapping [138]. Hence, they do not constitute a proper support for high quality graphene providing poor carrier mobility values. These findings emphasize the importance of the self-cleansing effect. In contrast, transition metal dichalcogenides (e.g. MoS_2, WS_2) are lipophilic in addition to their atomic flatness [275]. Accordingly, they form exceptionally clean and large common interfaces with graphene. The achieved carrier mobilities are high but still are surpassed by those obtained on hBN [138]. This might be explained by stronger scattering due to intrinsic crystalline defects in TMD as found by STM measurements [137]. Moreover, TMD are narrow bandgap semiconductors which limits their useable range as a gate dielectric [138].

Hexagonal boron nitride

If one could engineer an optimal substrate material for graphene, the result would be very close to hexagonal boron nitride (hBN). The individual layers of both materials share the hexagonal atomic structure having almost identical lattice constants[5] (hBN exhibits a 1.8% larger lattice constant [44]). The alternating B and N atoms form strong covalent sp^2 bonds which makes hBN chemically inert and almost free of charged impurities. Their density is found by STM measurements to be by 1-2 orders of magnitude lower in comparison to SiO_2 [120, 121]. Accordingly, the charge density fluctuations in graphene at the CNP are significantly smoothed (see Section 2.2.3 and Ref. 104). The surface of hBN is atomically flat and does not possess any

5 As discussed in Section 2.1.6, graphene's band structure remains unchanged by the hBN substrate as long as their crystallographic orientations are not well aligned to each other. Since the alignment requires special accurateness during device production, the misaligned, unperturbed case can be considered as the standard.

dangling bonds, which we will confirm by roughness measurements presented in the next paragraph. Hexagonal boron nitride also shows high affinity to graphene which results in large atomically clean interfaces when both materials are brought into contact [129, 138]. Possible contaminants are then pushed away and aggregate in few pockets, also referred to as bubbles. In addition, the phonon spectrum of hBN is favoring its use as a substrate. The optical phonon modes of hBN have twice the energy of the respective modes of SiO_2 which leads to a small thermal population even at elevated temperatures. The mechanism of remote-phonon scattering is negligible and causes a resistivity of only about 3Ω at room temperature [190]. This suggests that carrier scattering in graphene on hBN is dominated by graphene's intrinsic phonons (see Section 2.3.2). Hence, record room temperature mobilities above $100,000\,\mathrm{cm}^2/\mathrm{Vs}$ can be achieved [10].

As a big bonus, hBN is an insulator with a band gap of almost 6 eV [276]. This allows using it as dielectric in field effect gating. The dielectric breakdown field (perpendicular to the layers) is found to be about $12\,\mathrm{MV/cm}$ which is comparable to conventional SiO_2 due to the covalent bonding nature of hBN [277, 278]. Reported values for the out-of-plane static dielectric constant range from 3.5 to 5 [161, 278, 279]. In our devices we use a combination of SiO_2 with well-known thickness and exfoliated hBN flakes (whose height can be measured by AFM) as dielectric. Based on our experience, the dielectric constant of hBN can be estimated in good approximation to be the same as SiO_2 (3.9). The rather low dielectric constant is the only weak point in terms of its use for gating applications (however, this weakness is more than compensated by its excellent performance in other relevant parameters).

Thin sheets of hBN down to a single layer can be manufactured by the mechanical exfoliation technique, typically onto Si/SiO_2 wafer pieces [161, 280]. Depending on their thickness, graphene can become almost invisible when placed onto these sheets since the hBN flake offsets the total transparent dielectric thickness which is used to enhance the contrast of graphene by interference. However, an adjustment of the SiO_2 thickness would solve this minor issue. The bulk hBN crystals with high purity and typical dimensions of one millimeter are synthesized at high pressure and temperature ($4.5\,\mathrm{GPa}$, $1500°\mathrm{C}$) using barium boron nitride as solvent to suppress oxygen and carbon impurities [281, 282]. The crystal growth was pioneered by the National Institute for Materials Science in Japan which has established collaborations to most graphene research groups in the world providing their bulk crystals for mechanical cleavage. The small size and the crystal morphology of these hBN crystals, however, objects their straight application as bulk substrate. Thus, a challenging and still pending problem is the wafer-scale production of multilayered hBN

with homogeneous thickness[6] of at least 15-40 nm. Especially, these films should possess the same quality as exfoliated hBN flakes, e.g. in terms of intrinsic charged impurity density and surface flatness (see a recent review of this problem in Ref. 283). The achievement of this milestone would immediately pave the way towards wafer-scale high mobility graphene[7] for commercial applications.

Large areas of single-layer hBN are obtained by low-pressure CVD growth on metal foils (Cu [286, 287], Cu-Ni alloys [288], Fe [289], Pt [290, 291]) at about 1000°C using borazine as precursor. The latter is brought into the gas phase either directly by bubbling a carrier gas in liquid borazine or by moderate heating of ammonia borane powder which then decomposes into borazine. The so grown hBN monolayers show high crystalline quality and domain sizes of several ten to hundred micrometers. The self-limiting growth on metals can be overcome by working at higher (total and/or borazine partial) pressure, e.g. at ambient pressure. This results in several (ten) nanometer thick but polycrystalline (often nano-crystalline) and inhomogeneous hBN films [292–294]. Best results with respect to domain size have been achieved on polycrystalline Ni [294] or Fe [295] foils. However, the grown layer thickness is still found to be highly inhomogeneous, likely because of the polycrystalline foils which points to future improvements by using single-crystalline metal foils. Other studies have identified controlled surface segregation as an alternative way to obtain thicker hBN films with improved crystalline quality [296, 297]. The most promising approach, not requiring any post-growth transfer step, is the epitaxial CVD growth directly on dielectrics [293, 298, 299]. For example, it has been demonstrated that single-oriented multilayer hBN without any wrinkles can be grown on a sapphire wafer [293].

Large-scale all-CVD heterostructures are then produced by transferring graphene onto as-grown hBN and optionally capping it with another hBN layer [198, 275, 293, 295, 300–302]. The lack of crystalline quality of hBN and the necessary transfer process degrade the device performance. Especially, the transfer induces a high concentration of contaminants at the heterostructure interfaces. Hence, most reported mobility values are below $10,000\,\mathrm{cm}^2/\mathrm{Vs}$ and there are only few exceptions on hBN films possessing higher crystalline quality. The highest reported room temperature mobility reaches $24,000\,\mathrm{cm}^2/\mathrm{Vs}$ [295]. In addition to this slight mobility

6 As discussed earlier, this hBN thickness is typically required to ensure that hBN reliably compensates the surface roughness of the bulk wafer material and increases the distance to possible trapped charges in the underlying thicker substrate (e.g. SiO_2).

7 The quality of large area CVD grown single-crystalline graphene has already reached the level of mechanically exfoliated devices when it is transferred onto or encapsulated by exfoliated hBN flakes [203, 236, 284, 285].

improvement, the initial hole doping is reduced compared to SiO_2. Encapsulation further enhances the long-term device stability under ambient conditions [302].

Surface roughness measurements

We have measured the surface roughness of our bare SiO_2/Si substrates and several individual hBN flakes with different thicknesses (see Figure 3.2). All hBN flakes were freshly cleaved onto SiO_2/Si wafers without subjecting them to any further treatments, e.g. solvents or annealing steps. We investigated different SiO_2/Si wafer pieces purchased from various distributors. They were measured either as-received or after marker processing and standard wafer cleaning steps. No difference of statistical relevance could be attributed to a certain distributor or processing step. Non-contact mode atomic force microscopy was conducted using a commercial Park AFM NX10 in ambient conditions. This system is equipped with several anti-vibration measures, e.g. an acoustic enclosure and an active damping stage. The measurements presented here were performed while the AFM was located in our special Precision Laboratory building which additionally provided a comparatively low-noise environment. We determined the noise floor limit for our NC-AFM measurements to be equivalent to about 40 pm root mean square roughness (green line in Figure 3.2a). Since the enclosure of our AFM system is fitted with a temperature stabilization unit, thermal drifts are very small which allowed us to perform stable long-term scans with high resolution.

In the following we define the surface roughness as the root mean square value of the height distribution measured over an area of about two by two micrometer. This value is obtained directly using the data analysis software XEI that is part of the Park AFM system (after offset corrections and flattening). Based on 25 different wafer pieces, we determine the SiO_2 surface roughness to be 160 ± 17 pm in good agreement with reported values in the literature [120, 134, 303]. Our value corresponds to the typical interatomic distance between Si and O in amorphous SiO_2 which ranges between 160 pm and 180 pm [94, 304]. In contrast, other groups determine the SiO_2 roughness to be 250 pm by AFM [127, 161, 305], and 225 pm by STM measurements [104]. Some AFM measurements even indicate a much higher roughness of about 500 pm and above [124, 128]. We speculate that these differences likely result from different system calibration or data treatments. Further, modifications of the surface terminating groups might also alter the roughness, e.g. we find HMDS treated hydrophobic SiO_2 to possess a roughness of about 400 pm.

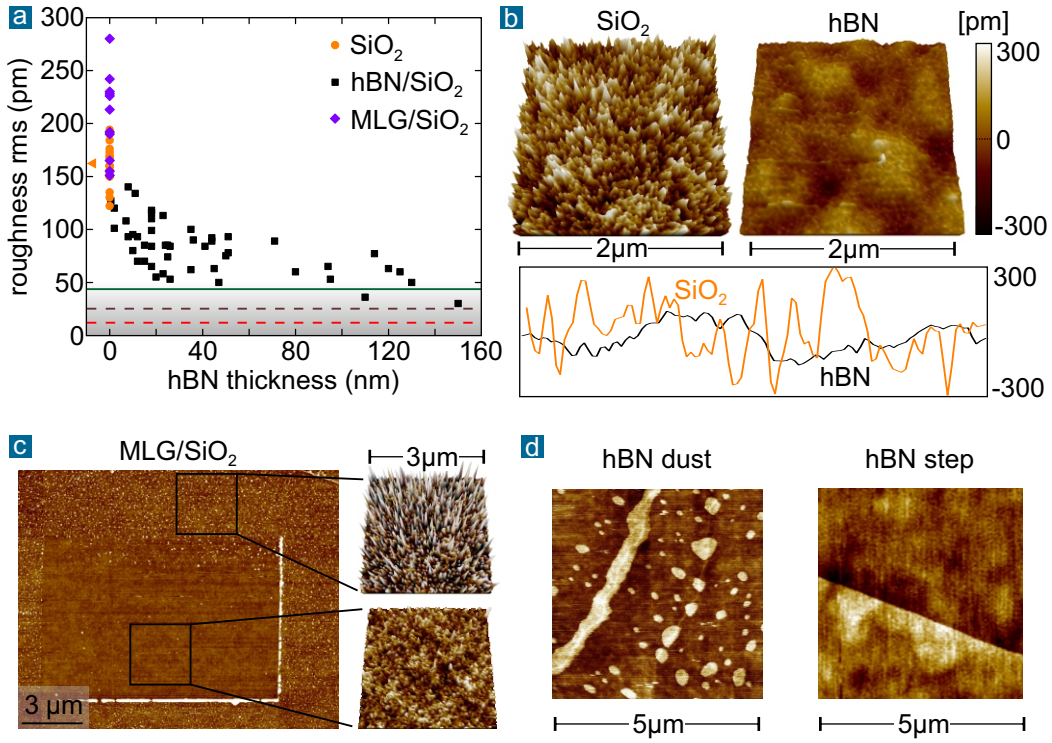

Figure 3.2. Non-contact AFM measurements: a: Root mean square roughness of a typical 2x2 micrometer area for a large ensemble of samples. The average roughness of SiO_2/Si substrates is 160 ± 17 pm (orange triangle). Roughness of exfoliated monolayer graphene on SiO_2 is often larger than the bare substrate. The roughness of hBN decreases with increasing flake thickness (each data point corresponds to an individual flake where we chose a dust-free and step-free area for determination of roughness). The green line marks the typical noise floor limit of the instrument. The brown dashed line corresponds to reported graphite and mica roughness [134], the red line to hBN-encapsulated monolayer graphene probed by TEM electron diffraction [125]. b: Exemplary surface maps of SiO_2 (rms 150 pm) and hBN substrate (25 nm thick, rms 62 pm). The lower panel shows two respective line cuts ($2\,\mu$m), which reveal only a slight corrugation of the hBN flake. c: Large area scan of a fully processed graphene on SiO_2 device where we employed the AFM based mechanical cleaning method to a certain (smaller) area in order to remove bulk processing residues which pile up at the tip reversal points (color scale as in b for zoom-ins and with ± 2 nm range for large area). In the cleaned area the roughness rms is 155 pm, while it is 500 pm in the non-treated reference area. d: Although hBN flakes are ideally flat, mechanical exfoliation can leave behind some hBN dust-like particles (or patches of thicker hBN) and cause steps on thicker hBN flakes. Color scale as in b, but ± 1 nm range for hBN dust image.

Very thin hBN flakes, alike monolayer graphene, conform to the SiO_2 substrate revealing similar roughness. With increasing thickness, however, the hBN flake more

and more compensates the wafer roughness which shrinks the observed value down to the hBN intrinsic value. Although we find the obtained values to show a notable spread, we clearly can identify this trend with increasing thickness in our measurements (see Figure 3.2a). As a rule of thumb, hBN flakes with a height larger than 15 nm fully compensate the initial SiO_2 roughness (see also Ref. 161). Two exemplary 3D maps and respective line scans of SiO_2 and a 25 nm thick hBN flake are shown in Figure 3.2b. It is obvious that the residual roughness of hBN mostly stems from a spatially extended weak corrugation while height changes for SiO_2 are stronger and happen on shorter length scales (here limited by our scan resolution). Note also the compelling similarity to electron hole puddles distributions observed by STM (see Figure 2.5a). It is intuitive that the spatially extended corrugation on hBN will not lead to substantial strain fluctuations which are a main cause of scattering.

With further increasing thickness we finally reach the noise floor of our instrument measuring values around 40 pm from which we can conclude that the roughness on hBN is at least decreased by a factor of four compared to SiO_2. We also obtain the same roughness for few layer graphene (graphitic) flakes which should be atomically flat, too, which indicates independently that we have reached the experimental limit of non-contact mode AFM in ambient conditions. Our findings agree with other AFM measurements revealing lower boundaries for hBN between 50 pm and 60 pm [303, 305], and for graphite and mica 23 pm and 36 pm, respectively [134] (brown dashed line in Figure 3.2a). In addition, STM groups have reported perfectly flat graphene on hBN on a nanometer length scale and an average roughness of 30 pm [104, 120]. Recently, TEM electron diffraction was used to determine the intrinsic roughness of graphene encapsulated between hBN [125]. While suspended (non-encapsulated) graphene shows intrinsic ripples [16], the latter are substantially smoothed by hBN. The study finds the (rms) roughness to drop from 114 pm (for suspended bare graphene) to only 12 pm when graphene is encapsulated by hBN (red dashed line in Figure 3.2a). This indicates not only the atomic flatness but also the absolute cleanness of the interface of such heterostructures. We emphasize that these record values are achieved on exfoliated flakes, while our own attempts to grow thicker hBN films by CVD result in roughness values above 500 pm (for a film thickness of about 5 nm that has been transferred onto SiO_2 using the polymer assisted transfer technique). This finding is sobering and highlights that growth and transfer routines have to be improved substantially.

The mechanical exfoliation method relies on the non-controllable cleavage process. We can neither predict the exact layer thickness nor is the latter always homogeneous. We illustrate this problem with two exemplary AFM scans shown in Figure 3.2d. On some flakes we find what we call in the following hBN dust-like particles and patches of slightly thicker areas. They cannot be removed by oxygen plasma or

thermal annealing so that we exclude them to be tape residues. Another problem are small height steps along the hBN flake. If such flakes are used to encapsulate graphene, both types of obstacles may act as pinning sites during the self-cleansing process and thus are highly detrimental to gain large-area ultra clean interfaces. As a consequence, we recommend to scrutinize the hBN flake surface prior to its use. For thinner hBN flakes below 40 nm, these faults are also detectable by optical microscopy, e.g. in dark field mode, which renders AFM measurements on each single flake unnecessary (note, the flake height can be also estimated by its color and transparency, see Figure 3.1). Since above 40 nm the optical contrast of step heights etc becomes too weak, our approach to use flakes between 15 nm and 40 nm is justified.

In contrast to early findings [127, 182], there is now consensus that graphene conforms with almost perfect fidelity to its substrate, e.g. as observed on mica [134], hBN [104, 161], and on SiO_2 [104, 128, 134, 136]. We have analyzed several exfoliated flakes onto SiO_2 (purple points in Figure 3.2a). The roughness is found on average to be larger than bare SiO_2 which contradicts other experiments that find graphene on SiO_2 as rough as the substrate or even slightly smoother [127, 128, 134]. A possible explanation might be that our flakes on pre-treated SiO_2 are additionally strained and wrinkled, e.g. by trapped molecules, or that graphene or the SiO_2 surface are charged. After device processing we even observe a roughness of about 500 pm due to PMMA residues (Figure 3.2c). However, when we brush away these contaminants by contact mode AFM (mechanical cleaning, see Section 3.1.2), the measured roughness in the cleaned area is 155 pm which corresponds to the bare SiO_2 value. Note, the mechanical cleaning does not remove or damage the graphene layer as monitored by Raman measurements (not shown). It seems that subsequent to exposure to resists and thermal curing of the latter, the graphene layer conforms much stronger to the SiO_2 roughness, in agreement with other reports. A corresponding, statistically significant change is also seen in our Raman studies (see Section 3.2.1).

3.1.4. Assembly of van der Waals heterostructures

In the previous sections we have discussed that polymer based handling and processing of bare graphene (and other 2D materials) will always deteriorate the interface (or surface) cleanness and thereby its electrical properties. Since these residues are impossible to remove completely and airborne contaminants further adsorb on uncovered surfaces, a straightforward solution is to avoid (or at least to minimize) them by encapsulation. Ideally, other van der Waals materials are used for this purpose.

In the following, any such arrangement of stacked layers of different van der Waals materials is referred to as van der Waals heterostructure or, in short, as stack. Let us consider the simplest heterostructure consisting of graphene encapsulated between a bottom and a top hBN flake. Using the aforementioned polymer assisted transfer of flakes would require two separate transfer processes: graphene onto bottom hBN, polymer removal by solvents, top hBN onto graphene/bottom hBN, and again polymer removal by solvents. During this bottom to top (layer-by-layer) assembling technique, graphene is still exposed to the polymer and solvents (wet transfer). Despite all cleaning attempts, the graphene-top hBN interface is contaminated by polymer residues which form dense bubbles in the whole area of the heterostructure (see details in Ref. 306).

Hence, it would be desirable to assemble these heterostructures by a dry method which does not expose the relevant interfaces to polymers and solvents such that these buried interfaces remain pristine. Near the end of 2013, Wang and coworkers demonstrated that it is possible to pick up a flake of a van der Waals material from the SiO_2 substrate just by means of another flake which is attached to a polymer stamp [10] (see Figure 3.3). The underlying key principle is universal. When two van der Waals materials are brought into direct contact, their mutual affinity and van der Waals forces can overcome the adhesion between the exfoliated (target) flake and the SiO_2 substrate. This assembling technique is thus referred to as pick-up or van der Waals stacking method. In this procedure, the top flake (e.g. the top hBN layer) is the only one in direct contact to the supporting stamp which guarantees polymer residue free interfaces between all other layers which are picked up in sequence until the whole stack is assembled. As last step, the latter is then released onto a bottom flake, typically hBN (on a SiO_2/Si substrate). Accordingly, the stack is built from top to bottom in contrast to the polymer assisted transfer of each individual layer which starts from the bottom.

The successful implementation of the pick-up method requires both a well designed setup and a smart choice of the polymer used as stamp (see Figure 3.3). Our stacking setup is basically an advancement of the setup we use for the polymer assisted transfer described in Section 3.1.1. The stamp itself is attached to a glass slide which is then mounted on a micromanipulator. This allows us to control the distance between the stamp and the substrate supporting the target flake. All flakes intended for the stack are freshly cleaved onto separate SiO_2/Si wafer pieces followed by the standard characterization methods. These wafer pieces are then fixed onto an aluminum holder whose temperature can be controlled. Its structure and dimensions are chosen such that an increase of one degree in temperature causes the aluminum block to expand by about one micrometer. This facilitates the task of bringing the flakes very slowly and gradually into contact. This slow thermal approach improves

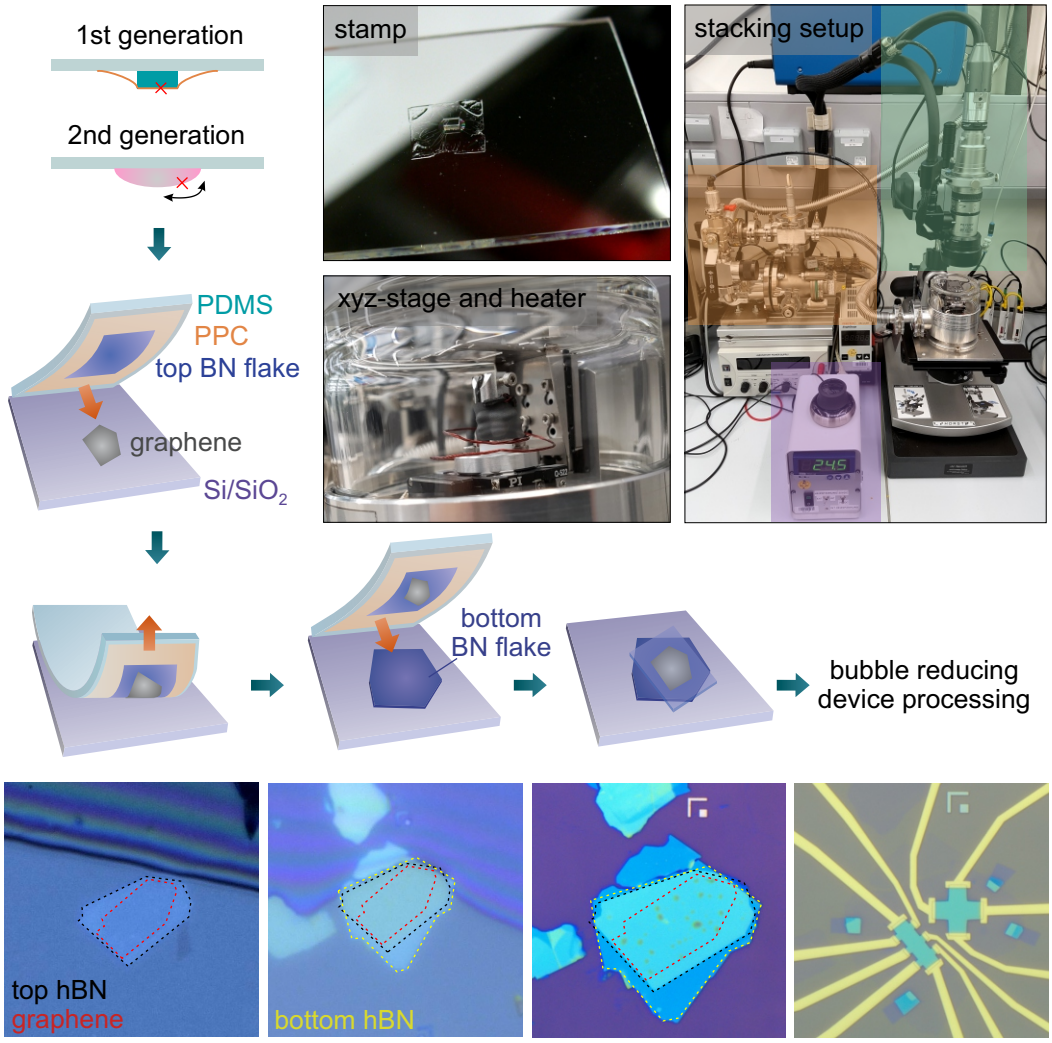

Figure 3.3. Van der Waals heterostructure assembly setup and stacking procedure. The photographs show the PDMS/PPC stamp on a glass slide and the home-built vacuum stacking setup. The latter consists of a long working distance zoom objective (shaded in green), temperature and position control (purple), a table-top turbo pump (orange), and an aluminum substrate holder with integrated heater mounted onto xyz-piezo stages (separate picture). The stacking procedure depicted in the drawings is explained in the main text. The bottom microscopic images were taken during the assembly procedure, while the right image shows the final device (marker size is 5x5μm, for reference).

the self-cleansing effect (see below). By means of two xy-micromanipulators, the target flake for the pick-up can be moved with respect to the stamp for proper alignment of the flakes. The stacking process is monitored by a long working distance

microscope equipped with a zoom lens. At first, we operated the setup in ambient conditions of a cleanroom environment (note, all devices presented in this thesis were produced in this way). Recently, we upgraded the system to perform the pick-up process in vacuum which together with pre-annealing of the target flake might reduce the airborne contaminants on the surfaces. Since some 2D materials react or degrade fast in ambient conditions, the stacking setup can also be placed into a glovebox with an inert gas atmosphere (Argon). This atmosphere, however, remains a source of hydrocarbons. The air unstable 2D materials can then be passivated by hBN flakes before they are removed from the glovebox for further device processing [307, 308].

The polymer stamp constitutes the other integral part of the stacking method. The basic idea is to exploit that the adhesion of a polymer film can be tuned by temperature. Its stickiness increases with temperature until the glass transition temperature is reached. Above the glass transition, the polymer loses its adhesion almost completely which releases the assembled stack from the polymer stamp in the final step. In addition, we have found that the yield of the pick-up process is increased at higher temperature, in agreement with independent studies [10, 244]. For the first generation of stamps, used to prepare all measured van der Waals heterostructures in this work, we followed the approach of Wang et al. (Ref. 10). As shown in Figure 3.3, a small, about one millimeter thick PDMS[8] stamp is attached onto an oxygen plasma treated microscope glass slide. In parallel, a one micrometer thick PPC[9] film is spincoated onto a bare silicon wafer piece. After baking at 80°C for 5 min, the PPC film is mechanically peeled off and transferred on top of the PDMS stamp on the glass side. Just before this transfer, the PDMS on the glass slide was briefly exposed to oxygen plasma in order to increase the adhesion between PPC and PDMS. The final PPC/PDMS stamp should then possess a smooth and particle-free surface. The stamp is typically slightly thicker at the edges of the PDMS such that they will touch the target substrate first during the approach. The ideal position for the stack assembly is thus in the center of the stamp, marked by the red cross in the drawing (Figure 3.3).

The stack assembly starts by picking up the top hBN flake with the bare PPC/PDMS stamp which is an easy task for flakes of at least several nanometer thickness. We

8 We use Sylgard 184 PDMS kit, mixing base:crosslinker 10:1. Vacuum/flush cycles are used to remove air bubbles after mixing. About 6-7g of the mixture is poured into a petridish with 12 cm diameter to get one millimeter thick PDMS. The PDMS is cured at room temperature for 2-3 days such that it remains flexible. A 5x5 mm piece is cut out and placed onto the glass slide. The final size of 1x1 mm is obtained by cutting with a razor blade. Check afterwards the PDMS surface to be flat and free of PDMS particles from the cutting process.

9 We use a solution of 12.5 weight percent polypropylene carbonate dissolved in anisole for spincoating. An approximately one micrometer thick film is achieved on a 8x8 mm sized bare silicon wafer piece at 1500 rpm for one minute.

preheat the substrate supporting the flake to about 40°C and slowly approach the stamp. Typically, one edge of the stamp will be the first to touch the wafer piece. We then stop the approach by the micromanipulator and instead increase the temperature in small steps up to $45 - 50$°C. Due to the thermal expansion of the aluminum block and the softening of the PPC film, the contact area between the substrate and the stamp slowly moves forward until it reaches the target flake (see microscope images shown in Figure 3.3). Ideally, the advance of this contact region should proceed very slowly[10] along the target flake. This can be achieved by slightly lifting the stamp by means of the micromanipulator, if necessary. As soon as the flake is completely in contact, the temperature is decreased again and the stamp is slowly lifted. The flake now adheres to the stamp. This procedure is then repeated for each flake of the planned stack using then the top hBN flake (or all already assembled layers) to pick-up the next flake. In the last step, the finished stack is released from the polymer stamp onto the bottom hBN flake by heating the substrate above the glass transition temperature of the polymer, in this case PPC ($T_g \approx 90$°C, release at about 120°C). The high temperature during the drop-down further helps to reduce bubbles and wrinkles in the stack (see next paragraph). Finally, the substrate with the stack is dipped into chloroform (followed by standard ACE/IPA wash), in order to remove PPC residues from the top hBN surface.

Bubble trouble

Looking at the resulting heterostructures fabricated by means of the above pickup procedure (bottom panel in Figure 3.3), one notices that they contain several bubbles which limits considerably the available area to fabricate clean devices. The formation of bubbles in van der Waals heterostructures occurs only between atomically flat van der Waals materials which exhibit a large affinity to each other [138]. It is thus a direct consequence of the self-cleansing effect. There is convincing evidence, e.g. by cross-sectional TEM on the buried interfaces, that these enclosures of several nanometer height contain trapped molecules, mostly hydrocarbons [129, 138]. The latter are airborne and cover the surfaces of 2D materials almost instantaneously after fresh mechanical cleavage [237]. As aforementioned, these contaminants are highly mobile on dangling bonds free, atomically flat van der Waals materials. During stacking, they are then squeezed into small isolated blisters leaving the rest of the interface atomically clean. The bubbles are typically rather dense and randomly distributed as seen in AFM and dark field optical microscopy (see Figure 3.4). Recent studies have examined the bubbles themselves and found that their typical size

10 Actually, for the first pick-up of the top hBN this is not critical. However, it is very important during the pick-up of further flakes to obtain properly self-cleaned interfaces.

and shape exhibits universal scaling determined by the competition between van der Waals attraction of the layers and their elastic deformation energy [309, 310]. Accordingly, bubbles give insight into the mechanical properties of the van der Waals materials and, even more interesting, they allow studying the trapped molecules which are experiencing pressures of up to one GPa inside the bubbles [310]. However, regarding electrical transport, the presence of bubbles increases the disorder in graphene and enlarges the initial amount of doping. Overall, the mobility decreases [311]. In particular, a shoulder on the electron side can emerge in field effect curves of the resistivity in case of high bubble density which indicates a substantial inhomogeneity in the channel. Hence, we have to avoid these bubbles in our sample designs which imposes additional constraints on achievable device dimensions.

Device processing from assembled van der Waals heterostructures

Device fabrication from encapsulated van der Waals heterostructures is conducted in a similar manner as for bare graphene on SiO_2. The final device geometry is protected with a 500 nm PMMA etching mask which is structured by means of electron beam lithography (EBL). The etching of the stack is realized in a fluorine based plasma. Initially, we used a mixture containing CHF_3 (fluoroform) and O_2 and operated the chamber in reactive ion etching mode (Diener Electronic). The homogeneous[11] etching results in a slope with an angle of about 45° (as depicted in Figure 3.4). This fortunate circumstance allows contacting the (likely one dimensional) edge of the buried graphene layer by thermal evaporation of metal covering the edges of the stack [10, 138]. Chromium (10 nm) is employed as adhesion layer covered by a thicker gold film (50-70 nm). The use of Chromium is critical to gain excellent electrical contacts. In the original work of Wang et al. [10], contact resistances below $500\,\Omega\,\mu$m were achieved over a broad density range (with a maximum of one $\mathrm{k}\Omega\,\mu$m at the CNP). The authors attribute these low values to the formation of a one-dimensional chromium-oxygen-carbon interface. Hence, it is important in this picture that the graphene edge is atomically clean before metal evaporation such that this interface forms unimpededly. Since the electrodes are patterned with EBL using an 800 nm thick PMMA double layer, PMMA residues might contaminate the edge of graphene. This can be avoided by either a short oxygen plasma treatment or an etch retreatment of the contact area in CHF_3/O_2 plasma prior to thermal evaporation of the metal layers [10, 43, 208].

11 The etch rate, however, slowed down considerably at large bubbles. This shows that hBN is etched fast but hydrocarbon species much slower.

In this work, the presented devices were etched and contacted following the above described procedure. We have found that the etching conditions using CHF_3 (in our Diener plasma chamber) and the obtained contact resistances vary among different samples. We speculate that this might be related to CHF_3 which can form thin passivation layers if the oxygen concentration is not high enough, e.g. due to poor control of the parameters by the instrument (compare to DRIE processes for silicon). We are now using a SF_6/Ar mixture and a more advanced and stable plasma source (Oxford ICP). Contact resistances have improved reaching the above reported values and they are achieved reliably (even without plasma retreatments prior to evaporation). This indicates the importance of purchasing dedicated machinery for 2D materials ensuring exceptional process cleanness (which is also true for thermal evaporators).

Scaling of the pick-up process and other 2D materials

The pick-up process can be repeated many times to build highly complex van der Waals heterostructures consisting of many layers and different 2D materials. The joint implementation of graphitic (i.e. multilayer graphene) back and top gates has become a standard for highest quality graphene devices when studying effects, e.g. quantum transport, at cryogenic temperatures (which requires at least seven layers in total [43]). Low contact resistances to several semiconducting 2D materials are obtained only when (few layer) graphene serves as contact material (which is then itself contacted by the metal edge contact) [312]. This requires even more repetitions of the stacking process. Finally, combining a vast variety of 2D materials is envisaged to create artificial layered crystals with novel and unique physical properties (reviewed in Refs. 15, 273, 274). This "atomic-scale Lego" approach has conceptual similarities to molecular beam epitaxy (MBE) but it offers compelling advantages. Van der Waals heterostructures are not limited by lattice mismatch issues and possess atomically sharp interfaces. Since the stack is built from separate bulk crystals, each layer has high crystalline quality and no interfacial mixing occurs in contrast to MBE. Recently, a Japanese research group presented an automated search and assembly system based on robot arms [313]. They demonstrated fast and reliable stacking of 30 layers. Hence, the fast prototyping of artificial layered crystals is becoming reality.

While the method seems to have no limits in vertical direction, there are severe limits regarding lateral dimensions. We still rely on exfoliation for the vast majority of van der Waals materials (especially, when requiring high crystalline quality). The flake size ultimately imposes limits to the total number of layers, e.g. with

respect to contacting several layers independently or regarding the risk of bubble formation. Recently, the pick-up technique has been adapted successfully to CVD grown graphene [236]. The authors could pick-up small single-crystals (about 100 μm) directly from the copper foil by means of a hBN flake and obtained similar mobilities than for exfoliated flakes. However, this method still relies on exfoliated hBN flakes and seems to work only when the copper foil below the graphene is oxidized (which is unlikely to happen for a larger graphene area on Cu). Hence, it remains unclear whether the pick-up assembly technique (and the inherent self-cleansing effect) can be scaled up to an industrial level.

3.1.5. Reliable improvement of van der Waals heterostructures

We have identified two measures that reliably enhance the quality of van der Waals heterostructures subsequent to their assembly: thermal annealing and mechanical ironing by means of contact mode AFM. They are described in the following paragraphs. We briefly discuss our present working hypothesis about how the improvements come about. Statistical evidence is provided in the following section covering electrical transport measurements and Raman spectroscopy. Moreover, we present a more comprehensive study in our dedicated publication (Ref. 183) where results from this work are combined with other low temperature transport measurements conducted in our group. Both demonstrate unambiguously the substantial quality gain achieved by following our proposed routine.

Step 1: Thermal facilitation of the self-cleansing effect after stack assembly

Instead of accepting the constraints in device dimensions imposed by the bubbles, one may ask whether it is possible to reduce the amount of bubbles after the stacking procedure or minimize their formation during the stacking process? The general answer is that a higher temperature apparently speeds up the diffusion of (trapped) hydrocarbons and so the self-cleansing effect. When we anneal our assembled van der Waals heterostructures above 300°C, the many small bubbles merge into few larger ones (see dark field images in Figure 3.4). At 500°C the bubbles disappear to a large extent (likely they burst[12]) and instead we observe several tens of nanometer high wrinkles, often along the edges of the graphene flake. This might be related

12 It has been shown that when scratched by a sharp AFM tip, the bubble bursts and the trapped molecules leak from the bubble until the remaining van der Waals interaction seals the blister again [129].

to the observations by other groups that entire graphene flakes can slide and rotate on hBN flakes (for several micrometers) at such temperatures, under the premise of thoroughly clean interfaces [47, 314] (see also Section 2.1.6 discussing the emergence of Moiré superlattices). When encapsulated by hBN, graphene shows negligible environmental sensitivity and remains unaffected by high temperature annealing and even by plasma treatments [161, 315, 316]. Hence, thermal treatment of the stacks at high enough temperatures (around 500°C) significantly increases the available area for device designs *and* results in excellent electrical properties since it reliably increases the cleanness of the interfaces (as shown in Section 3.2.2 and sketched in Figure 3.4). That means, the density of charged impurities is substantially reduced. As discussed in more detail in the following, we find the 500°C annealing step subsequent to the stack assembly to be crucial for the effectiveness of our proposed second step involving a mechanical treatment after device processing. In other words, the charged impurity density in close proximity to graphene, at the hBN-graphene interfaces, first has to be lowered such that strain fluctuations turn into the dominant remaining source of disorder. The mechanical treatment then attempts to release these leftover strain fluctuations.

Having demonstrated that annealing after stack assembly is beneficial, a higher temperature during the pick-up process is useful in two ways. It increases the yield of the pickup process *and* the self-cleansing effect is already enhanced during the pick-up itself. In order to be able to work at higher temperature, the stamp design has to be adapted by switching to other polymers with higher glass transition temperatures. Alternative candidates are among others polycarbonate (PC) [317], GelPak/bare PDMS [318], PMMA [138], or Elvacite acrylic films [319]. Among these, Elvacite and PC have the highest glass transition temperature of about 160°C. For the second generation of stamps we use a drop of Elvacite dissolved in anisole which is placed on the glass slide and baked at 200°C. The ideal spot for the stack assembly is located away from the center of the stamp hemisphere since the latter will first touch the target substrate (see Figure 3.3). As a side effect, the hemisphere shape improves the process of the slow movement of the polymer-substrate contact area (acting like a roll-off stamp). The above described pick-up procedure remains the same except that the flakes are picked up around 90°C and the final stack is released at 190°C. By this optimized process, the amount of bubbles is reduced significantly and sometimes we[13] find them to be even completely absent. Furthermore, the pick-up process has higher yield and even works for flakes that adhere strongly to the SiO_2 surface, e.g. due to plasma pre-treatment of the substrate prior to exfoliation.

13 The employment of Elvacite stacks was tested by Dr. Youngwook Kim in our group. As-fabricated stacks show much larger bubble-free areas and exhibit improved electronic quality. Those most recent devices could not be measured and analyzed anymore as part of this work.

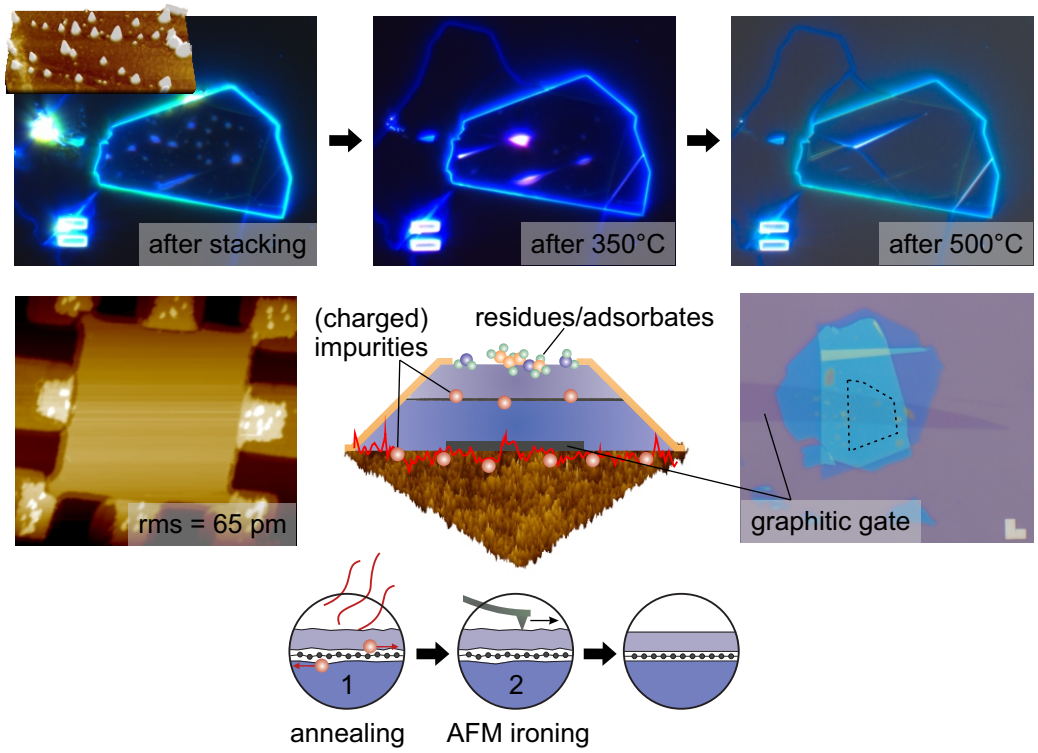

Figure 3.4. Optimization steps before and after device processing of heterostructures. Top panel shows optical dark field images taken directly after the stack assembly and after thermal annealing at different temperatures in a forming gas (Ar/H_2) atmosphere (5x5μm marker size). A typical AFM scan of a non-annealed area is shown in the inset (10x20μm and $\pm$2nm color scale). During annealing the numerous small bubbles aggregate to larger ones and at 500°C they often vanish completely. The significantly increased bubble-free area allows designing larger devices. The key drivers for quality improvement are sketched in the bottom center drawing (compare also with Figure 2.5 in Section 2.2.3 summarizing all possible sources of extrinsic disorder). Inherent to all stacks, the bottom hBN compensates the roughness of SiO_2 and the top hBN flakes protects graphene from direct contact to residues and adsorbates. The thermal annealing at 500°C after stacking (1) enhances the self-cleansing effect and residues, i.e. charged impurities, are removed efficiently from the graphene-hBN interface. The mechanical treatment by means of contact mode AFM after the device processing (2) releases strain fluctuations which constitute the remaining source of extrinsic disorder in annealed, high quality devices. Remote charged impurities located at the SiO_2 surface can be screened by a graphitic bottom layer that serves as a back gate (example shown at the bottom right, dashed line indicates hardly visible graphene). An exemplary AFM scan of a 3x3μm device that was subjected to our proposed routine is depicted at the bottom left. The image reveals only a very small roughness. Electrical contacts to graphene are established by evaporating metallic side contacts after the entire heterostructure stack has been etched with a slope of about 45°.

In Section 3.1.1, we have discussed that such treatment increases on average the dimensions of exfoliated 2D materials, at the expense of a stronger interaction with the substrate. We find that graphene layers on plasma pre-treated SiO_2 cannot be picked up at 40°C (using PPC) but can be reliably lifted at about 90°C (using Elvacite). Hence, stacks of larger dimensions can be obtained. Similar results have been recently reported elsewhere using a modified PPC stacking routine [244][14]. These authors find that thicker flakes (e.g. the top hBN) can always be picked up (independent of temperature and substrate pre-treatment). This flake is dropped onto graphene at 110°C by lifting the PPC/PDMS stamp. The temperature is then reduced to 40°C and the PPC/PDMS stamp is lowered to pick-up both the top hBN and graphene since both flakes now adhere stronger to each other than graphene to the substrate. The hot contact of the flakes at 110°C is found to be crucial to obtain completely bubble-free areas which supports our theory that the self-cleansing mechanism is thermally accelerated. The authors also show that the adhesion forces between the fused flakes are strong enough to tear the graphene flake. If the top hBN flake is smaller than the graphene flake, the uncovered part of graphene is left behind (a cutter-like mechanism). In addition to these optimizations, we have equipped our stacking setup with a vacuum chamber (see Figure 3.3). This allows us to anneal the target flakes in vacuum prior to pick-up in order to minimize hydrocarbon contamination (i.e. charged impurities). This further improves the interface cleanness and reduces the probability for bubble formation. However, we emphasize that, despite all these measures to optimize the pick-up process itself, the post-assembly annealing step at 500°C is still necessary to obtain reliably high quality devices.

Step 2: AFM ironing after device fabrication

After stack assembly and ideally thermal annealing at high temperature, i.e. 500°C in Ar/H_2, we fabricate devices from those areas of the heterostructures that are considered free from any bubbles based on optical microscopy and non-contact AFM. Following the device processing steps described in detail in the preceding Section 3.1.4, this involves at least two times the application, baking, and removal of PMMA resists as well as evaporation of metals and lift-off processes. A second thermal treatment at 350°C in Ar/H_2 is typically performed on the finished devices in order to reduce the residues on the surface of the top hBN layer. However, we note that

14 After this chapter was written, another group (Ref. 320) found that the removal of bubbles is possible when the fully assembled stack is released slowly onto the target substrate at 180°C. This again supports our picture that contaminants become highly mobile at elevated temperatures.

this second annealing step is not critical as discussed further below. We also found a higher temperature than 350°C to deteriorate the electrical contact to graphene.

Electrical transport measurements are finally performed on these fabricated devices. Despite the use of the van der Waals assembly technique and the encapsulation with hBN for all devices, a substantial variation persists, as will be discussed in detail in Section 3.2.2. While the aforementioned high temperature annealing step already improves the quality and reduces the variation, we have identified a second, post device processing step that reliably boosts the quality further, increases substantially the overall yield of excellent devices[15], and even allows us to heal initially poorly performing devices. This step, we will refer to as AFM ironing, involves a mechanical treatment of the heterostructure stack's surface by an AFM tip, as described in detail in the following.

In analogy to the mechanical cleaning of uncovered graphene devices on SiO_2 from "bulk" residues by means of contact mode AFM (as described in Section 3.1.2 and shown in Figure 3.2), we have tested a similar mechanical treatment on our hBN-encapsulated devices. The AFM tip is operated in contact mode and scanned across the entire active channel of the device[16]. This active area has been preselected based on the absence of bubbles in AFM non-contact mode images. We note the important difference that the tip is in mechanical contact to the surface of the top hBN layer but not to the graphene layer as it is in case of uncovered devices on SiO_2. We will demonstrate via electrical transport studies in Section 3.2.2, both statistically when comparing a large set of devices as well as a single device pre- and post-treatment, that this additional mechanical treatment after all device processing steps improves reliably the device quality and disorder level. This is reflected in higher mobility and smaller residual carrier density values. Moreover, we find that the aforementioned high temperature annealing step at 500°C, conducted after stack assembly but before device processing, is an important prerequisite for the effectiveness of this mechanical treatment of the final device (see also Section 3.2.2).

At first, it is tempting to attribute the observed quality improvement to the same mechanism as for AFM treated uncovered graphene samples, even though the achieved

15 Otherwise, obtaining such high quality devices is a matter of statistics. Typically, a multitude of devices has to be fabricated and those with only mediocre quality are usually not considered for further measurements. A procedure not only followed by us but also by the research community in general. To clarify, our devices do not exceed the reported mobility values by other groups on "cherry-picked" devices, but we achieve a much higher yield of such high quality devices.

16 We use a commercial AFM system (NX-10 from Park Systems) and AFM tips that are normally intended for non-contact mode operation (Park PPP-NCHR). They have a tip radius of curvature around 10 nm. Tip force in contact mode was set between 50 and 150 nN. More details can be found in our publication (see Ref. 183).

improvement in devices that are not encapsulated is much smaller. After all, the AFM tip is expected to clean the top hBN surface from residues which may serve as a source of remote charged impurities. However, we have collected evidence that this picture of the quality improvement mechanism is incorrect. In the following we summarize the main findings published in our work dedicated to this topic (Ref. 183). It covers both the electrical transport measurements presented in Section 3.2.2 as well as low temperature quantum oscillation studies which are not subject of this work. They support our hypothesis that the main beneficial effect does not arise from the removal of remote charged impurities on the top hBN surface but from ironing away nanometer scale corrugations at the interfaces between hBN and graphene (as sketched in Figure 3.4). Due to the thermal enhancement of the self-cleansing effect the charged impurity density close to the graphene layer, that means at the interface, is already minimized and random strain fluctuations ultimately prevail. The AFM procedure then releases these strain fluctuations even further.

First of all, we often do not observe any pile-up of "bulk" residues at the boundaries of the scan area after conducting the AFM treatment - in contrast to the case of uncovered graphene on SiO_2 (see Figure 3.2). This is checked by performing non-contact AFM measurements before and after the treatment. The 350°C annealing step after device fabrication seems to remove "bulk" residues quite efficiently from the top hBN surface. Moreover, it can be shown that the device properties remain the same after intentionally applying and removing another polymer layer on top of the heterostructure, i.e. the high quality achieved by the AFM treatment is maintained even though the top hBN surface is dirty again (see our work in Ref. 183, and Ref. 10). This implies that encapsulated graphene is not sensitive to residues on the top hBN surface. Finally, our technique also improves devices equipped with a graphitic top gate that screens the potential of remote charged impurities located on the top hBN surface [183]. Removing such impurities should then have no impact when the picture of remote charged impurities as main source of disorder holds.

Our hypothesis that an ironing of the interface roughness takes place is backed by the work of Couto and coauthors in Ref. 18. The authors have provided substantial evidence, i.a. by studying quantum interference effects, that random strain fluctuations constitute the predominant disorder source in high quality graphene devices (see our work in Ref. 183 and Section 2.3.2 for a detailed discussion). In the low temperature limit, they observe a correlation between the mobility and residual carrier density that can only be reproduced in a disorder model involving random strain fluctuations. As shown in panel d of Figure 3.8 in Section 3.2.2, the values obtained on our devices, that were subject to the proposed annealing and AFM ironing routine, follow this predicted correlation between mobility and residual density nicely.

Since the quality of our devices is notably better, the range for which the theoretical model can be tested is even enlarged, and still the model holds. In our work (Ref. 183), we also demonstrate that even larger, AFM detectable bubbles, which are known to consist of trapped residues at the interfaces, can be manipulated, i.e. moved around, by contact mode AFM. The larger the bubble the more force is necessary (up to a certain size limit). This indicates that it should be possible to flatten nanometer scale roughness at the interface with comparatively small forces[17]. Very recently, after our results were published, two other studies appeared that provide further experimental evidence for our hypothesis that random strain fluctuations, i.e. microscopic out-of-plane corrugations, are the main source of disorder in high quality, hBN-encapsulated devices [185] and that they can be manipulated after stacking and device fabrication [184]. In Ref. 184 it was shown that microscopic corrugations can be reduced by applying small uniaxial strain to the graphene-hBN heterostructure. This results in an increase in mobility while the residual carrier density at the CNP decreases which can be explained nicely by the random strain fluctuations model.

In conclusion, our proposed device treatment based on the combination of two straightforward steps, thermal annealing at high temperature (500°C) after assembly and AFM ironing after device processing, improves reliably the quality of van der Waals heterostructures. This procedure is likely only applicable to 2D materials with a high mutual affinity such that self-cleansing occurs. We have verified the positive impact of our method also on hBN-encapsulated MoS_2 devices (see our work in Ref. 183). The thermal treatment enhances the self-cleansing such that charged impurities are efficiently removed from the interfaces. What remains are random strain fluctuations which are addressed by AFM ironing. The disorder level after AFM treatment is on average substantially lower compared to untreated van der Waals heterostructures.

Further optimization: Gate designs without involving the SiO_2/hBN interface

In Section 2.2.3 we have presented and discussed recent experimental findings that remote charged impurities can contribute to the density fluctuations in already very clean devices. In addition, a gate voltage hysteresis[18] of the resistance arises

17 Obviously, the necessary force should be also a function of the top hBN layer thickness. However, for the typical thickness on the order of 10-20 nm used for encapsulation applying moderate force values works properly.

18 Hysteresis here refers to the different gate voltage values at which the CNP peak in the resistance is observed when comparing up and down sweeps of the gate voltage. The hysteresis persists

also for hBN-encapsulated graphene when a silicon back gate is used [321]. That means when the electrical field lines penetrate through the SiO_2 dielectric and the respective SiO_2-bottom hBN interface (see Figure 2.5h). Both are known to host trap charges and dangling bonds saturated by polar molecules which are supposed to be the culprits of the hysteresis [84, 93, 94, 240, 321]. A similar hysteresis effect occurs when a mica flake is used as gate dielectric, but vanishes as soon as hBN is used as (the only) gate dielectric in the same heterostructure [322]. Furthermore, laser irradiation leads to local doping of the graphene layer due to photoinduced charge transfer to charge traps in dielectrics [291, 324, 325]. We have observed such doping in early experiments performing electrical and Raman measurements in parallel, too (not subject to this thesis). The CNP of the channel area is pinned to the gate voltage which was applied during laser exposure. We find the doping to slowly decline within hours which can be accelerated by annealing at higher temperature.

To conclude, hysteresis effects can be substantially reduced by using only hBN as dielectric and not involving the SiO_2/hBN interface [321]. Furthermore, a graphitic back gate shields the graphene layer from remote charged impurities of the SiO_2/hBN interface. These remote charged impurities otherwise affect quantum transport (oscillations) through inhomogeneous broadening. For this reason, thin graphitic (few layer graphene) gates now often replace the substrate-embedded silicon back gate in recent designs of van der Waals heterostructures (for example in Refs. 43, 47, 206, 326, 327, sometimes also thin evaporated metal gates are used alternatively). They can be realized in top and/or back gate configuration. The use of graphitic gates, however, is more beneficial when conducting (quantum) transport measurements at low temperatures, while their positive effect on room temperature properties, which are mainly covered in this thesis, is smaller in our experience than the effect of the above described post-assembly optimization steps. Hence, most of our studied devices were not equipped additionally with graphitic gates. Figure 3.4 shows an exemplary stack of hBN-encapsulated graphene which has been placed onto a few layer graphene flake serving as back gate in the future device. The risk of bubble formation increases with each additional flake in the stack, ultimately limiting the device size. Hence, such complex heterostructures with a sufficient bubble-free area are much easier to realize using Elvacite and the high temperature pick-up procedure.

also at cryogenic temperatures and its amplitude increases with larger voltage scan range and sweep speed (see Refs. 321–323, and own experience). This hysteresis effect is often not shown in the literature (typically only one sweep direction is included). However, hysteresis can be a critical effect for graphene applications.

3.2. Sample statistics

In this section we present and analyze statistical findings based on many graphene samples we have measured in the course of this thesis. This allows us to understand how graphene's key electronic parameters are influenced by certain processing and cleaning steps, e.g. thermal annealing and self-cleansing effects. Our analysis will further emphasize and fortify the substantial improvement of the properties by hBN encapsulation in comparison to bare graphene on SiO_2. We start by discussing the main insights gained from Raman spectroscopy studies at room temperature. The latter not only serves as a powerful tool to distinguish the number of graphene layers but also reveals valuable information on strain and doping. The method is fast and non-invasive which allows monitoring graphene's electronic quality during device fabrication. The second half of the section presents the main results obtained from electrical transport measurements (here in absence of a magnetic field, while the next chapter covers in detail the magneto-transport in classically weak magnetic fields). We investigate the difference in initial doping level and gate voltage hysteresis between hBN encapsulated and uncovered graphene. Finally, mobility values and the residual carrier densities at the CNP will be compared. We demonstrate that Raman spectroscopy reaches its limits in terms of distinguishing between different grades of high quality devices while electrical transport measurements are more sensitive and precise in assessing the highest quality. Statistical evidence for the beneficial effect and reliability of our proposed post-processing method to improve van der Waals heterostructures, which we introduced at the end of the preceding section, is provided. The main findings of this section were also published in Ref. 183.

3.2.1. Pre-assessment of graphene's quality by Raman spectroscopy

Raman spectroscopy of graphene in a nutshell

Raman spectroscopy is one of the most important tools to examine graphitic materials (sp^2-bonded carbon allotropes), e.g. carbon nanotubes and graphene [328]. Its application to study graphene started soon after graphene's experimental availability and gathered large momentum since then [329, 330]. In this section we will not provide a general introduction to Raman spectroscopy and neither in detail to all effects and non-prominent peaks in graphene's spectrum for which we refer to excellent reviews [19, 328, 330, 331].

The Raman spectrum of graphene is characterized by its simplicity possessing only two (three) prominent peaks in case of pristine (slightly defective) lattices [331]. The G-band (G for graphite) appears at 1582 cm^{-1}, in the undoped and unstrained case, and is associated with a normal first order Raman (intra-valley) scattering process. It involves the double degenerate (iTO and LO) phonon mode at the Brillouin zone center which corresponds to the in-plane vibration of both sublattices in opposite directions (E_{2g} symmetry)[19]. In contrast, the D and 2D peaks are related to second-order Raman (inter-valley) scattering processes. They are located at about 1350 cm^{-1} and 2700 cm^{-1}. Their exact positions depend on laser excitation energy, since they undergo a blue-shift with increasing energy [330, 332]. The Raman scattering process involves two iTO phonons near the Dirac point for the 2D peak or one iTO phonon and one defect for the D peak [331]. Hence, the latter is only present when graphene's lattice exhibits a substantial amount of defects (or sp^3-bonds). Its absence constitutes a criterion for high crystalline quality and a low level of intrinsic disorder. The 2D peak, however, is always active by double resonance, and is strongly dispersive with laser energy due to a Kohn anomaly at the K point [330]. Note, the 2D peak is sometimes also called G' peak to emphasize that it is always present (in particular for other carbon materials such as CNT [328]).

The inherent double resonance process makes the 2D band highly sensitive to details of the electronic band structure [329, 331]. Accordingly, its shape (and width) evolves with increasing layer thickness (see Figure 3.5a). Monolayer graphene exhibits one single and narrow peak due to its single band (Dirac cone). For bilayer, this peak breaks up into four sub-peaks related to four different double resonance processes which arise from additional bands in its band structure. These four peaks have very similar energies such that we measure only their envelope. As the layer number increases, the band structure transforms finally into that of graphite which is reflected by its characteristic 2D band shape. In addition, the relative intensity of the G peak increases steadily with layer thickness. Hence, mono- and bilayer graphene can be unambiguously identified by their Raman fingerprint[20].

Besides its use to identify the layer number, Raman spectroscopy has become a versatile tool to probe graphene's electronic and mechanical properties [330]. The exact peak positions $\omega_{G,2D}$ and their full width half maxima (FWHM) $\Gamma_{G,2D}$ depend

19 In full analogy, hexagonal boron nitride possesses only one characteristic Raman peak located at 1366 cm^{-1} which is also associated with its E_{2g}-symmetric phonon mode [280].

20 The identification of few layer graphene above three layers, however, remains very challenging due to the absence of unique changes in the 2D peak shape. Recent studies on low wavenumber Raman spectroscopy below 100 cm^{-1} have found that few layer graphene (and other 2D materials) can be also characterized by shear modes around 40 cm^{-1} [333]. Furthermore, twisted multilayer graphene possesses layer breathing modes which reflect the twist angle [334].

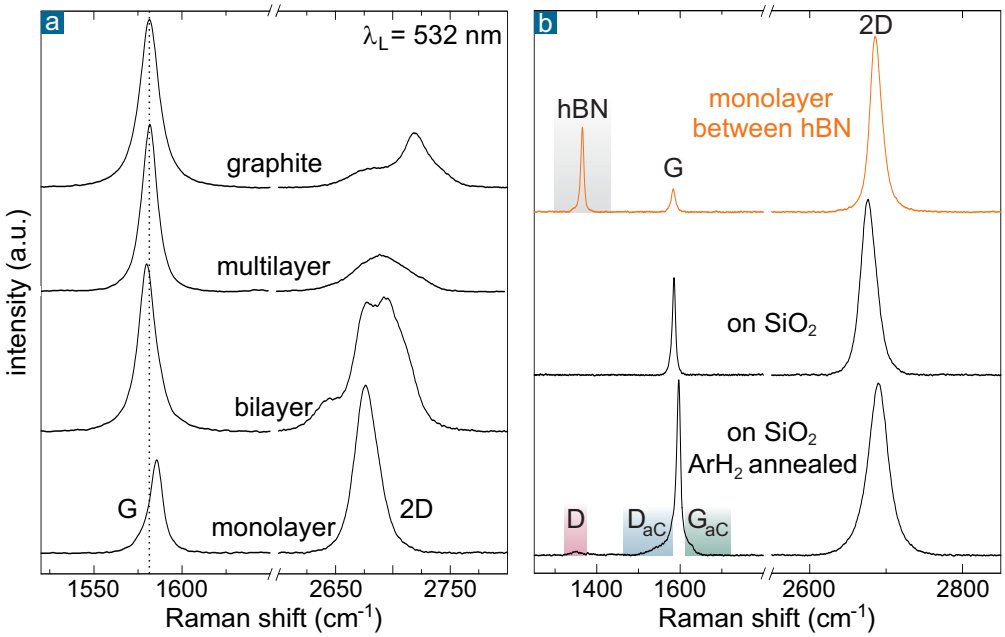

Figure 3.5. Exemplary Raman spectra of monolayer graphene and thicker layers under ambient conditions. a: Evolution of the Raman spectra with layer thickness (exfoliated on SiO_2). The shape of the 2D peak and the intensity ratio between the peaks constitute a clear fingerprint to identify mono- and bilayer graphene. The vertical dashed line corresponds to the pristine G peak position (about $1582\,cm^{-1}$), indicating a deviation for mono- and bilayer graphene due to strain and/or doping. b: Comparison between hBN encapsulated, bare graphene on SiO_2, and graphene on SiO_2 after thermal treatment at 350°C in Ar/H_2 atmosphere. For graphene between hBN, the peak height ratio is the largest and the 2D peak width is the smallest. In addition, the hBN related peak at $1366\,cm^{-1}$ is observed. After annealing SiO_2 supported graphene and re-exposing it to air, the G peak blueshifts and the 2D peak becomes broader in comparison to untreated graphene on SiO_2. Often we observe a small D peak (related to defects) and side bands around the G peak which we labeled D/G_{aC} following the nomenclature of Ref. 264. They are likely related to amorphous carbon deposited on the graphene flake during annealing (which may stem from burnt-off tape residues in the vicinity of the flake).

on temperature [335–337] and doping [86, 338, 339], and give insight into both global strain [340–342] and nanometer-scale strain variations [123]. The main challenge is to interpret the Raman spectra correctly by disentangling the interrelated effects caused by temperature, strain and doping. The parameter space can be shrunk by performing all measurements at room temperature and by defining a standard laser wavelength to avoid the dispersion of the 2D band. Hence, we use a green laser with a wavelength of 532 nm ($E_L = 2.33\,eV$) to achieve conformity with the most important

reports in literature (which use either exactly the same wavelength [123, 124] or 514.5 nm, a very similar one [265, 343]). We further adapt our data analysis method to these standards and conduct our measurements in a Raman system offering high precision and reproducibility[21].

Separation of doping and strain effects

Several experiments based on Raman mapping of individual flakes have contributed to disentangle the two remaining parameters, mechanical strain and charge doping. Their effect can be separated by a vector decomposition model which is represented in a diagram spanned by the peak positions of the G and 2D band (see Figure 3.6a) [124, 265]. The origin (red point) of the ω_{G}-ω_{2D} space is determined by the intrinsic frequencies of graphene not subjected to any strain or doping. This point can be estimated by measuring Raman spectra of pristine suspended graphene (e.g. taking the average of a map) [124, 265, 332]. In our diagram, we take the origin $\omega_{\mathrm{G}}^{0} = 1581.6 \pm 1\, cm^{-1}$ and $\omega_{\mathrm{2D}}^{0} = 2678.6 \pm 1\,\mathrm{cm}^{-1}$ as determined in Ref. 124 for the same laser wavelength.

Global strain applied to the undoped graphene lattice shifts the frequencies of both peaks by a fixed ratio. The relative slope for uniaxial strain $(\Delta\omega_{\mathrm{2D}}/\Delta\omega_{\mathrm{G}})_{\epsilon,\mathrm{uni}}$ is about 2.2 at which this value can vary slightly depending on crystallographic directions or when biaxial components contribute [265, 341, 342]. The black dashed line in our diagram (Figure 3.6a) indicates the effect of strain isolated from doping effects. Above (below) the origin, the flake experiences compressive (tensile) strain, respectively. Furthermore, strain fluctuations on the nanometer-scale (substantially smaller than the spot size) lead to broadening of the peak width for both lines [123]. For weak strain fluctuations and low (native) doping level (e.g. on hBN), the strain fluctuations inherent broadening of the G line can be only observed when electronic broadening effects (see below) are suppressed by a strong magnetic field [123]. The electronic broadening of the G peak is also suppressed by a larger (native) doping level in absence of a magnetic field. Then, strain variation related contributions

21 Our Raman measurements are conducted by means of a confocal system (Bruker Senterra) in the ambient conditions of a cleanroom. The spot size is below one micrometer. The instrument is automatically and continuously calibrated through parallel measurement of a neon spectrum which ensures a wave number precision and drift stability below $0.1\,\mathrm{cm}^{-1}$ (Bruker SureCAL). The spectral resolution is better than $3\,\mathrm{cm}^{-1}$ and the whole spectrum is taken in three steps such that the step size of the acquired data is $0.5\,\mathrm{cm}^{-1}$. After subtraction of the background, we fit single Lorentzians to each peak of the spectrum to obtain the exact position and FWHM (in accordance with the data processing in the literature).

prevail, e.g. in case of (CVD grown) graphene on SiO_2, and the G peak broadening is clearly related to regions dominated by bubbles and wrinkles [123].

The FWHM of the 2D peak, in contrast, is hardly influenced by doping (even high carrier densities result in a negligible broadening of only three wave numbers [123, 332]). The 2D peak width is, however, substantially widened when the amplitude of the strain fluctuations increases (compared to the strain-free value of $17.9\,\mathrm{cm}^{-1}$ [123]). Hence, the analysis of the 2D band width can be used to determine the strength of strain variations in absence of a magnetic field and independent from the (native) doping level. Since these fluctuations have been identified as a major contribution to scattering [18], the measurement of Γ_{2D} reveals valuable information about the expected electrical properties of graphene on different substrates [123, 124]. The experiments of Ref. 123 have demonstrated that the origin of the peak broadening is unambiguously related to these strain fluctuations on the nanometer scale because the slope between Γ_G and Γ_{2D} has the same value (2.2) as $\Delta\omega_{2D}/\Delta\omega_G$ for global strain (when electronic contributions to Γ_G are suppressed). The laser spot size of 500 nm up to one micrometer then averages over a range of slightly shifted G or 2D peaks. This results in the broadening.

Due to non-adiabatic electron-phonon coupling (Kohn anomaly at the zone center), the G peak shifts to larger wave numbers for both electron and hole doping [86, 338, 339, 344]. This stiffening is much more pronounced for ω_G than for ω_{2D}. Moreover, the sign of the doping determines whether ω_{2D} increases or decreases [86, 265, 344]. Hole doping leads to an almost linear shift of both peaks, similar to the effect of strain but with a different slope $(\Delta\omega_{2D}/\Delta\omega_G)_{hole} \approx 0.75$ (blue dotted line in Figure 3.6a) [265]. In contrast, $\Delta\omega_{2D}/\Delta\omega_G$ becomes non-linear in case of electron doping resulting in a bent trace in the ω_G-ω_{2D} space (indicated as gray dashed line in the vector schematic) [124, 344]. In the vector decomposition model each measured combination of ω_G and ω_{2D} is then projected onto the hole doping and strain axes [124, 265]. This approach is supported by electrical measurements which typically reveal an initial hole doping for graphene on SiO_2 and low doping on hBN (see Section 3.2.2). Hence, the vector model is valid for these substrates[22].

As mentioned before, the width of the G peak is strongly influenced by the carrier density [86, 123, 339]. The FWHM Γ_G has a maximum of about $15\,\mathrm{cm}^{-1}$ at zero

22 However, when graphene on novel substrates with unknown native doping is measured, one needs to consider also electron doping as a possible contribution to the ω_G/ω_{2D} position. The diagram still allows categorizing the substrate effect regarding strain and doping. If significant doping is concluded, increased Coulomb scattering will occur and the substrate is less suitable for graphene [124]. The sign of the doping can then be determined by means of electrical transport measurements.

(average) doping and sharpens with increasing density (or doping). This sharpening arises due to Pauli blocking of the phonon decay channel into electron-hole pairs [86, 344]. Furthermore, the peak intensity ratio I_{2D}/I_G decreases with increasing carrier density [86, 338, 344]. In summary, doping has a remarkable impact on graphene's Raman spectrum and in particular on the G peak. Unfortunately, the type of doping remains obscure.

Following the vector decomposition model, the ω_G-ω_{2D} space is split into three areas (A1-A3, see Figure 3.6a). In the first area (A1) the graphene is (hole) doped and experiences global tensile strain, whereas in the second area (A2) the flake is compressed and also doped. The third area (A3) is in principle forbidden since both doping types result in blue-shifts of ω_G. However, we observe that hBN-encapsulated graphene shows (ω_G,ω_{2D}) positions within this area, in agreement to other reports [124, 343]. It has been observed that only the 2D peak frequency is shifted to higher wave numbers, while the G peak frequency remains constant, when graphene is placed onto hBN, and even shifted further when it is encapsulated between hBN [345]. Similar findings are also reported for graphene on transition metal dichalcogenides [124]. Hence, these small deviations to the left of the black dashed line (corresponding to ω_{2D} shifts) do not indicate any detrimental effect of the electronic properties of graphene (which are known to be excellent on these substrates). Three different theories have been proposed to explain this shift into the "forbidden" area A3. Detailed calculations predict a logarithmic phonon anomaly leading to a small kink of the blue trace [265]. Another theory invokes the high sensitivity of the 2D peak frequency to the Fermi velocity which can change on hBN due to van der Waals interactions and renormalization (arrow labeled Δv_F in Figure 3.6a). And finally, the enhanced dielectric screening by hBN might tune the Kohn anomaly and cause the 2D peak shift (arrow labeled screening) [124, 345]. Fortunately, the effect is weak and does not notably alter the conclusions drawn from the vector decomposition model.

Large ensemble statistics on the influence of substrates and processing steps

The hitherto published experiments study in each case only one or few individual samples [123, 124, 265, 343]. These samples are scanned by the laser to obtain spatially resolved Raman maps. The studies find the peak positions and full width at half maxima to vary along each sample within a certain range (cloud-like scattering of the data points similar to Figure 3.6a). These variations reflect the spatially inhomogeneous doping and, in particular, strain distribution within the sample. The analysis then focuses on the collective shifts of these scattered data points due

to thermal annealing [265, 343], or substrate related effects (mainly the difference between SiO_2 and hBN [123, 343]). However, none of these studies presents measurements of a large ensemble of samples to provide information about sample to sample variation. Moreover, the reported measurements are performed on pristine flakes or CVD graphene films which are not processed into devices (e.g. spincoating and removal of PMMA masks for etching and metal evaporation). Additionally, no systematic comparison between different ways to exfoliate or handle graphene is made: Ref. 265 uses uncovered, exfoliated flakes on SiO_2 which was pre-cleaned with a piranha solution. Ref. 343 places CVD grown graphene onto a thin hBN flake and a reference SiO_2 area by wet transfer during which the graphene remains uncovered. In contrast, Refs. 123, 124 pick-up exfoliated [123] or CVD grown graphene [124] with a hBN flake in a dry process and release them on the substrate whose effect is to be tested, e.g. a second hBN flake (completing the stack) or SiO_2 (and other substrates [124]). This approach implies that the surface of graphene is always covered by hBN. Hence, it allows separating the substrate effect from airborne and resist/polymer contaminants. However, it does not represent the typical, uncovered (large area) graphene device on SiO_2 since the size of the top hBN flake is limited.

The purpose of our analysis is to complement the so far reported variations within each sample, which we refer to as "intra-sample" variations in the following, by an investigation of the inter-sample spread. We further extend our study to different stages of the device processing, both for uncovered and encapsulated graphene. Finally, a third aspect is considered: We examine how treatments of the SiO_2 substrate prior to exfoliation influence the doping and strain level (see also Section 3.1.1). Our main findings are summarized in Figure 3.6. Each data point is based on the Raman spectrum of an individual graphene sample (or device) made from exfoliated flakes. We emphasize that taking into account only one spectrum per sample neglects possible intra-sample variations. Hence, in our case, each data point is a random representative of a cluster of data points which is measured along a sample. However, despite the neglected intra-sample variations, we observe a clear clustering of the data points (for similarly treated samples). This makes the following conclusions exceptionally substantiated.

Our analysis comprises graphene samples on three different substrates. First, we consider graphene flakes exfoliated directly onto SiO_2 substrates with different pre-treatments (described in Section 3.1.1). We distinguish between completely untreated substrates, which have never been in contact with polymer resists (as-received), and substrates equipped with markers, which were cleaned by solvents and oxygen plasma after marker processing by EBL. The exfoliation onto as-received wafer pieces (blue squares) has the lowest yield but the exfoliated flakes can be picked up already at 40°C during the stacking process (for which we use these flakes

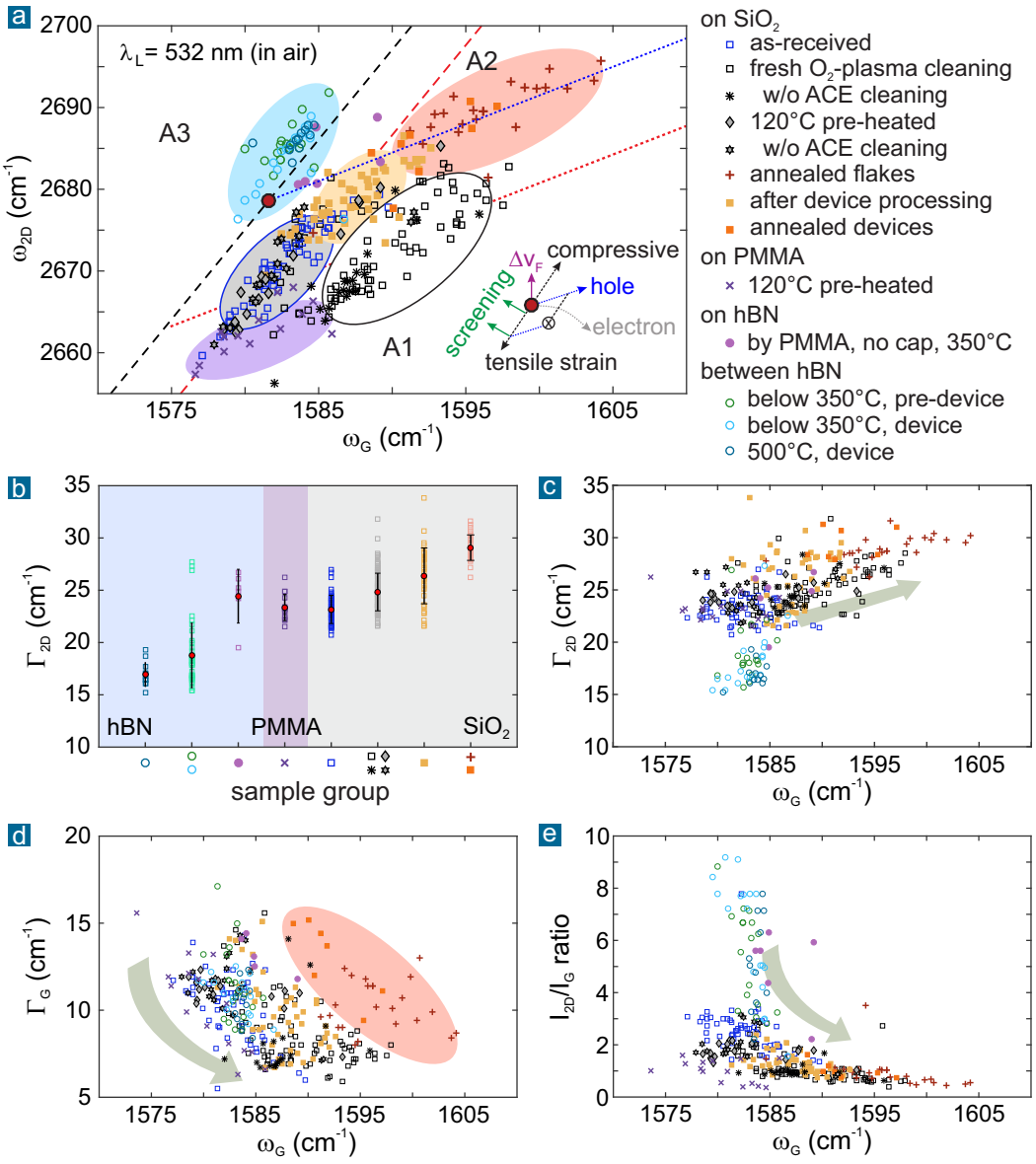

Figure 3.6. Statistics based on the evaluation of Raman spectra of a large ensemble of graphene samples. Each scatter point corresponds to an individual sample. Different processing steps and substrates notably alter the values of $\omega_{G,2D}$ and the FWHM $\Gamma_{G,2D}$ of these two prominent peaks (see discussion in main text).

exclusively). In contrast, for substrates with markers, we increase the yield by pretreating the SiO_2 surface in a twofold manner (see Section 3.1.1). Either we refresh

the plasma treatment directly before exfoliation (black symbols) or we heat the substrate for several minutes at 120°C in air and exfoliate during the cool-down (gray symbols). In the latter case, we used only wafer pieces whose last plasma treatment was at least one day old (such that it showed the same water contact angle than as-received SiO_2). In order to remove tape residues from the SiO_2 substrates (and to test the adhesion of the flakes), the substrates with the exfoliated flakes are typically washed by ACE/IPA dips before further processing. We also account for effects which might arise from this cleaning step by performing reference measurements on flakes without acetone cleaning (star-shaped black and gray symbols, respectively). We then fabricate devices from these exfoliated flakes on SiO_2 substrates with markers (yellow squares). This exposes the graphene to PMMA resists and solvents, and thermal curing of the resists at 160°C (see Section 3.1.1 for details). Some of these devices were further annealed in Ar/H_2 atmosphere at 350°C (orange squares). As reference, we also annealed exfoliated flakes which did not undergo the device processing (red crosses). In both cases, after device processing and/or annealing, we find no difference between the two pre-treatment methods of the SiO_2 substrate. Accordingly, we have merged these data points in Figure 3.6.

The second type of substrate is a PMMA layer (on $PVA/SiO_2/Si$) onto which we exfoliate graphene after the substrate has been heated to 120°C in order to increase the yield (purple x marks). These flakes are then transferred onto hBN flakes, as third type of substrate, using the polymer-assisted process (see Section 3.1.1). Subsequent to the standard device production steps (electron beam lithography and plasma etching), the graphene on hBN devices are cleaned in ACE/IPA and annealed in Ar/H_2 atmosphere at 350°C (purple circles). Note that graphene's surface is uncovered in this device design ("wet transfer"). Furthermore, we perform Raman measurements on hBN encapsulated flakes and the respective devices (open circles). They are produced by means of the van der Waals pick-up method for which graphene is never exposed to polymers ("dry transfer", see Section 3.1.4). We distinguish measurements on stacks or fabricated devices which were annealed at a maximum temperature of 350°C (or below[23]) and devices which were processed following our current procedure (500°C annealing after stacking, and 350°C after device processing).

Discussion of the Raman statistics presented in Figure 3.6

Different SiO_2 pre-treatment steps obviously lead to different starting points regarding strain and doping level of the exfoliated flakes (see Figure 3.6a). Graphene on

23 The very first stacks were only annealed at 200°C on a hot plate before device processing.

as-received or pre-heated SiO_2 substrates, that means without fresh plasma treatment, exhibit a moderate hole doping level with small spread (gray oval area with blue boundary). Most data points are located along the red dashed line which indicates a rather homogeneous doping level but varying tensile strain between the samples. The range of these sample to sample variations in terms of strain is comparable to reported variations along an individual sample [123, 265]. In contrast, fresh oxygen plasma treatment of the SiO_2 surface leads to a substantial increase of the overall hole doping (and its spread) while the strain level and its spread do not change considerably compared to the case without plasma cleaning (dotted red line). As discussed in the previous sections, this can be understood by the removal of hydrocarbons which are normally attached to the polar silanol groups. The bare silanol groups then are in vicinity to the graphene lattice which leads to charge transfer and increases the chemical reactivity. Accordingly, acetone and IPA treatment of the graphene flakes on plasma treated SiO_2 enhances the hole doping and the scattering between the samples becomes larger (compare black stars and squares, respectively). A similar trend is not observed on SiO_2 without fresh plasma treatment which demonstrates the connection between silanol groups and reactivity.

After device processing (which involves application, thermal curing, and removal of PMMA resists), the initial tensile strain is partially released while the moderate[24] hole doping level persists (yellow oval area). In particular, also the devices made from flakes which were exfoliated onto plasma treated SiO_2 show similar ω_G, ω_{2D} coordinates which implies a reduction of the hole doping. On average, we find both pre-treatments to result in devices with similar strain and doping level. Hence, we argue that during device processing hydrocarbons (or other molecules) might diffuse between graphene and SiO_2 and attach to the silanol groups (similar to the observed reduction of the hole doping after immersion into water [232]). However, if graphene flakes or devices on SiO_2 are annealed in Ar/H_2 atmosphere at 350°C, a strong hole doping arises (red oval area in Figure 3.6a). While the emergence of strong hole doping after annealing is in agreement with other reports [95, 232, 265], we find our samples to be almost free of strain after the thermal treatment, in contrast to reports about strong compressive strain [232, 265]. In the above theoretical framework, the annealing process is understood to clean the interface between graphene and SiO_2 such that an efficient electron transfer to the silanol groups can happen [95, 232]. Since graphene conforms stronger to the SiO_2 roughness after annealing [95], the interface is better sealed and the hole doping becomes almost irreversible [232]. This

24 The absolute hole density is on the order of $10^{12}\,cm^{-2}$, as determined by electrical measurements (see Section 3.2.2).

is in particular true for devices after fabrication since annealing is typically the last process step before electrical measurements.

We find the hole doping effect after annealing to occur only on SiO_2 but not on hBN (see also Refs. 95, 343), in accordance with the silanol group model. Graphene on hBN devices (filled purple circles) are fabricated from graphene flakes which have been exfoliated onto PMMA (purple oval area). The moderate doping and the substantial tensile strain, initially experienced by graphene flakes on PMMA, decrease notably after their wet transfer onto hBN followed by annealing at 350°C in Ar/H_2 atmosphere. Similarly, the doping and the strain of graphene flakes on as-received SiO_2 wafer pieces (blue squares) almost entirely vanish when the flakes are encapsulated with hBN by means of the dry pick-up method described in Section 3.1.4 (all open circles and turquoise area). Not even annealing at 500°C induces any doping. If at all, it causes weak compressive strain (which is, however, difficult to extract from the ω_G, ω_{2D} coordinates due to the general shift of the 2D position on hBN). Graphene encapsulated by hBN and annealed at 500°C exhibits the smallest variation of ω_G, ω_{2D} compared to all studied samples. This proves the good reproducibility of the device quality using our proposed annealing routine (see Section 3.1.5). As discussed in detail in the supporting information of our publication (Ref. 183), the spread in the Raman data points (scattering range) remains the same after performing additionally the AFM ironing on the 500°C annealed devices (hence not shown additionally in the figure), despite the strong improvement in the mobility and residual density values we observe in electrical transport measurements. This indicates that Raman spectroscopy is not able to distinguish between different grades of high quality devices, while electrical transport measurements are inherently more sensitive to the achieved improvements.

Our findings are further supported by analyzing the full width half maxima of the Raman peaks and their intensity ratio (Figure 3.6b-e). The width of the 2D peak varies by up to $10\,cm^{-1}$ between different substrates and processing steps (Figure 3.6b). For uncovered samples on SiO_2, the corresponding strain fluctuations on the nanometer scale increase slightly with the number of processing steps (on average $\Gamma_{2D} \approx 25\,cm^{-1}$). In particular, thermal treatment at 350°C leads to a significant increase of the broadening (on average $\Gamma_{2D} \approx 30\,cm^{-1}$). This finding supports the above picture that graphene conforms better to the substrate after annealing and hence to the roughness of SiO_2 in this case (and likely the observed release of the global strain happens at the expense of local strain fluctuations). Figure 3.6c shows the 2D peak broadening with respect to the measured G peak position. The latter is also influenced by global strain. If one presumes the validity of the vector decomposition model, the strain-free G peak position only reflects the hole doping value. This value can be obtained by projecting each data point onto the blue dotted line

in Figure 3.6a (along a vector parallel to the black dashed line). This would result in a shift to larger ω_G values for most data points, and in particular for the measurements of bare (non-processed) flakes on SiO_2. Then, the trend between ω_G^{hole} and Γ_{2D} for uncovered samples on SiO_2 becomes even more obvious (as indicated by the arrow). The connection between the strength of the strain fluctuations (on the nanometer scale) and the hole doping supports the above picture of the doping mechanism on SiO_2 (strain increases reactivity to adsorbates and residues that act as dopants).
For hBN encapsulated graphene samples we find $\Gamma_{2D} \approx 18\,cm^{-1}$ which corresponds to the intrinsic broadening (Figure 3.6b). However, this value is only reliably achieved when the stacks are annealed at 500°C after assembly in order to obtain atomically clean interfaces (following our proposed routine). Again, the effect of the AFM ironing cannot be clearly distinguished by comparing Γ_{2D} values, but unambiguously by studying electrical transport (see next subsection). The importance of (self-)cleaned interfaces becomes evident when we consider the much larger peak broadening ($\Gamma_{2D} \approx 24\,cm^{-1}$) for non-encapsulated graphene which was placed on a hBN flake by the wet transfer process (similar values for wet transfer of CVD graphene onto hBN are reported in Ref. 236). Although the graphene on hBN samples were annealed at 350°C, the strain fluctuations are on the same order than on SiO_2. Hence, the elevated values of Γ_{2D} are most likely related to contaminants from the transfer process which locally strain the graphene lattice [124].

As discussed earlier in this section, doping leads not only to a blue-shift of the G peak frequency but also to a reduction of the G peak width and to a smaller I_{2D}/I_G ratio. We observe these trends in our measurements as shown in Figure 3.6 where we plot Γ_G and I_{2D}/I_G over ω_G. In contrast to experiments which control the carrier density of a single graphene device with a fixed amount of strain by gating [86, 339, 344], the doping in our case varies statistically (in uncontrolled fashion) between different samples with different strain level. Hence, we have to account for possible strain contributions to the G peak positions of the data points (as discussed above). Considering this aspect, it is clear that the correlation between ω_G and Γ_G (or I_{2D}/I_G) on the doping level is subject to stronger scattering of the data points. For all samples but the annealed ones on SiO_2, we find the same trend between ω_G and Γ_G as observed in Ref. 338 studying a larger ensemble of samples. Most of the annealed samples on SiO_2 (red area in Figure 3.6d), however, possess a larger G peak width than one would expect for the corresponding large hole doping (or ω_G). We attribute this broadening to the contribution of strain fluctuations (on the nanometer scale) which we know to be the largest for these annealed samples (based on the Γ_{2D} values in figure panel b). As discussed in Ref. 123, the averaging over strain fluctuations can dominate the G peak broadening in case of low quality samples with significant doping (when electronic broadening is negligible).

Since the I_{2D}/I_G ratio is not influenced by strain fluctuations, the expected correlation with ω_G is more pronounced (Figure 3.6e). In particular, when the strain contributions to ω_G were taken into account which would result in larger effective frequencies ω_G^{hole} (e.g. for the blue squares, gray symbols, and purple x-marks). Furthermore, the intensity ratio on hBN is enhanced by an interference effect. Since hBN exhibits a similar dielectric constant as SiO_2, the effective oxide thickness increases by about the hBN flake thickness which varies between 15 nm and 50 nm from sample to sample. According to Ref. 346, the I_{2D}/I_G ratio is then enhanced by interference and can exceed the value of undoped graphene samples on SiO_2 with 290 nm thickness by up to a factor two (depending on the hBN thickness). This explains the different I_{2D}/I_G ratios between the graphene samples on hBN which are typically very low doped and should exhibit similar I_{2D}/I_G ratios.

3.2.2. Electrical transport - device benchmarks and intrinsic limits

In contrast to Raman spectroscopy, electrical transport measurements allow us to quantify precisely the initial doping level and its sign as well as the residual charge carrier density at the CNP and the mobility. As described in Section 2.2.1, we apply a voltage V_g to the (back)gate in order to tune the average carrier density in graphene devices, e.g. shaped into a Hall bar geometry, while monitoring the longitudinal resistance. From the latter, we then calculate the resistivity ρ (or the conductivity $\sigma = 1/\rho$) taking into account the sample geometry. Examples of such field effect curves are shown in Figure 3.7. The gate voltage is converted into the gate controlled average density n_g using Equation 2.18. The voltage value V_{CNP} corresponding to the charge neutrality point is defined as the gate voltage value of the resistivity maximum ρ_{max} while the gate factor β is known either from independent Hall measurements or is calculated based on the gate geometry and the properties of the dielectric. As discussed in detail in Section 2.2.3, the total carrier density n_t in the graphene sample is only equal to the gate controlled average density n_g away from the CNP (that means for large $V_g - V_{CNP}$ or n_g values, corresponding to the single carrier type regime). In the vicinity of the CNP, however, both electrons and holes are present when $n_g = 0$ (due to thermal activation and disorder). Hence, the system is always conductive. We have introduced this residual total carrier density at the CNP $n_{t,0}$ in Sections 2.2.3 and 2.3.3 when discussing the two carrier type Drude model which presumes the system to be spatially homogeneous. Within the channel approximation and by combining Equations 2.30 and 2.56, the conductivity

in the absence of a magnetic field is given by Equation 2.57 which we repeat here for convenience:

$$\sigma = \frac{1}{\rho} = \mathrm{e}\,\mu\,\mathrm{n_t} = \mathrm{e}\,\mu\,\sqrt{\mathrm{n_{t,0}^2} + \mathrm{n_g^2}}\,. \tag{3.1}$$

Note that we also have assumed that the electron and hole mobilities are equal. Within this framework, one can obtain several important electrical parameters of a graphene device just by analyzing the field effect curves of the resistivity (even without magneto-transport measurements when the gate factor is known). As discussed in Section 2.3.3, the residual carrier density at the CNP $\mathrm{n_{t,0}}$ can be estimated from a double-logarithmic diagram of the conductivity versus the gate controlled average density $\mathrm{n_g}$. An extrapolation of the linear trend of $\log\sigma$ at higher densities down to the saturated value of the conductivity at low density, σ_{min}, yields a convenient estimate of $\mathrm{n_{t,0}}$. An example of this procedure is shown in Figure 3.8a. Since the resistivity maximum ρ_{max} is known, an estimate of the mobility at the CNP, μ_{CNP}, can be obtained with the help of Equation 3.1 for $\mathrm{n_g} = 0$ ($\sigma_{\mathrm{min}} = 1/\rho_{\mathrm{max}} = \mathrm{e}\,\mathrm{n_{t,0}}\,\mu_{\mathrm{CNP}}$). Furthermore, at high densities, e.g. $\mathrm{n_g} = 10^{12}\,\mathrm{cm^{-2}}$, the mobility in the single carrier type regime can be deduced within the Drude model from the measured resistivity (corresponding to μ_{FE} in Equation 2.38).

In the following paragraphs we present and discuss the statistics of the aforementioned electrical parameters (residual density and mobility) obtained from a large set of graphene devices. As in the previous Raman part, we distinguish between two different sample types: bare graphene on SiO_2 and hBN-encapsulated graphene. In addition, the samples are classified according to the processing steps they underwent. We demonstrate that van der Waals heterostructures can be reliably improved following our proposed post-assembly treatment involving a high-temperature annealing and a mechanical ironing step (see Section 3.1.5). Intrinsic limits set by phonon scattering and thermal activation of charge carriers are reached at room temperature. Electrical transport measurements conducted on a multitude of devices allow us to provide clear evidence for the beneficial effect when combining these two post-assembly steps, both statistically across a large set of devices as well as when comparing one and the same device before and after the treatment. It is also shown that the high-temperature annealing step is a crucial pre-requisite to achieve the improvement upon AFM ironing. In contrast to electrical transport measurements, Raman studies do not offer this capability since they do not reveal a statistically relevant difference among mechanically treated and untreated hBN-encapsulated graphene samples, as discussed in the preceding subsection. Our final discussion focuses on the properties at room temperature which is of importance for future applications of graphene. We establish easy-to-measure criteria at room

temperature, e.g. in terms of the resistivity values at the CNP and at large densities, devices of excellent electronic quality have to meet.

Initial doping level and hysteresis effects in field effect gating

All bare devices on SiO_2 and most hBN-encapsulated graphene devices are gated by means of a silicon back gate that is integrated into our standard Si/SiO_2 substrates. The gate dielectric therefore consists of either only 290 nm thick SiO_2 (resulting in a gate factor of $\beta = 7.4 \cdot 10^{10}\,\mathrm{cm^2/V}$) or an additional, bottom hBN flake with a thickness that varies among the samples (reducing the gate factor since the EOT increases). If a hBN layer is present, we determine the gate factor through an independent Hall measurement (see Chapter 4). Some stacks are also equipped with a thin graphitic back gate. Then, only the bottom hBN layer serves as gate dielectric. The gate factor is then much larger and on the order of $\beta = 10^{12}\,\mathrm{cm^2/V}$. The applied voltage range was chosen such that the resistance peak at charge neutrality (for $V_g = V_{CNP}$) can be unambiguously identified, although we do attempt to stay within the save window of the gate dielectric (see Section 2.2.1 for details)[25]. The gate voltage is swept upwards (i.e. the voltage is increasing) and downwards (i.e. the voltage is decreasing) at a speed of $0.1 - 0.5\,\mathrm{V/s}$ or $0.01 - 0.05\,\mathrm{V/s}$ when using a graphitic gate.

A comparison of the up- and down-sweep of the field effect curve reveals hysteresis. In particular, bare graphene devices on SiO_2 exhibit a significant shift of the resistivity maximum position ($V_{CNP}^{up/down}$) and, in extreme cases, even its amplitude changes (see Figure 3.7a,b). Although this hysteresis effect is frequently neglected in basic research (and, for example, only the up or down sweep is shown/taken into account), it is of great importance for real-world applications of graphene demanding excellent reproducibility. In order to compare the strength of the hysteresis and the initial doping level among different samples with different gate factors, we define the initial doping density as

$$n_i = -0.5\,\beta\,(V_{CNP}^{up} + V_{CNP}^{down})\,, \tag{3.2}$$

and the hysteresis effect with respect to the density as

$$n_{hys} = \beta\,|V_{CNP}^{up} - V_{CNP}^{down}|\,. \tag{3.3}$$

25 For graphene on SiO_2 devices which were annealed at 350°C in forming gas (Ar/H_2 atmosphere) and, in general, for graphene devices exposed to air during measurements, higher voltages up to 75V have to be applied. For vacuum baked (non-annealed) devices on SiO_2 and for the stacks we typically stay below 30V. If a graphitic back gate is employed and the gate factor increases, this voltage span is scaled down accordingly (typically, by a factor of ten).

Here, $V_{CNP}^{up/down}$ are the positions of the resistivity maximum for an up- and down-sweep. Positive n_i values correspond to initial electron doping and negative values to initial hole doping.

Our study distinguishes[26] between graphene on SiO_2 devices which have been annealed in 350°C Ar/H_2 atmosphere prior to electrical measurements and non-annealed devices. These are the devices that according to the Raman spectrum fall into the red and yellow area in Figure 3.6a, respectively. We find that the uncovered graphene devices are extensively influenced by the environment (main results are presented in Figure 3.7 and Table 3.1). The samples exhibit a substantial initial hole doping and strong hysteresis when operated in air. Both properties remain present also in vacuum but their level decreases. The hole doping can be further lowered after baking the devices at 140°C in vacuum for a few hours. In some cases, we even observe a change to electron doping. The hole doping, however, recovers when the chamber is vented again with (humid) air. The vacuum baking step is found to be ineffective on 350°C annealed devices which in general possess a higher extrinsic hole doping (compare the red data points of annealed and vacuum baked devices and the blue data points of non-annealed and vacuum baked devices in Figure 3.7c, and Table 3.1). This observation is in full agreement with the conclusions drawn in Section 3.2.1 based on our Raman measurements and studies published elsewhere [95, 253, 265]. The hole doping is explained by charge transfer to volatile adsorbates which are stabilized by the stronger interaction between graphene and dangling bonds on the SiO_2 surface after thermal annealing. Hence, we recommend to skip the 350°C annealing treatment for bare graphene on SiO_2 devices. In the following chapters we focus on non-annealed samples when referring to graphene on SiO_2 and all measurements are performed in vacuum after baking the samples at 140°C. As a side effect, this step also improves slightly the mobility and reduces the disorder induced residual carrier density at the CNP (compare the resistivity peak broadening in Figure 3.7b).

Although we can make clear statements about bare graphene on SiO_2 devices based on the ensemble statistics, the spread between the samples is (too) large to predict the characteristics of a single device. We even observe substantial variations between devices produced in the same batch with exactly the same processing steps. While in basic research one can select the best suited ones, the sample to sample variations (and environmental dependence) constitute a huge obstacle for applications of graphene. Ideally, the initial doping should be fixed (to zero) and hysteresis effects

26 Device fabrication is obviously a pre-requisite for electrical transport measurements. Hence, we are not able to differentiate between the two different pre-treatment methods of the SiO_2 substrate prior to exfoliation since they lead to very similar results *after* device processing as already shown in the Raman section.

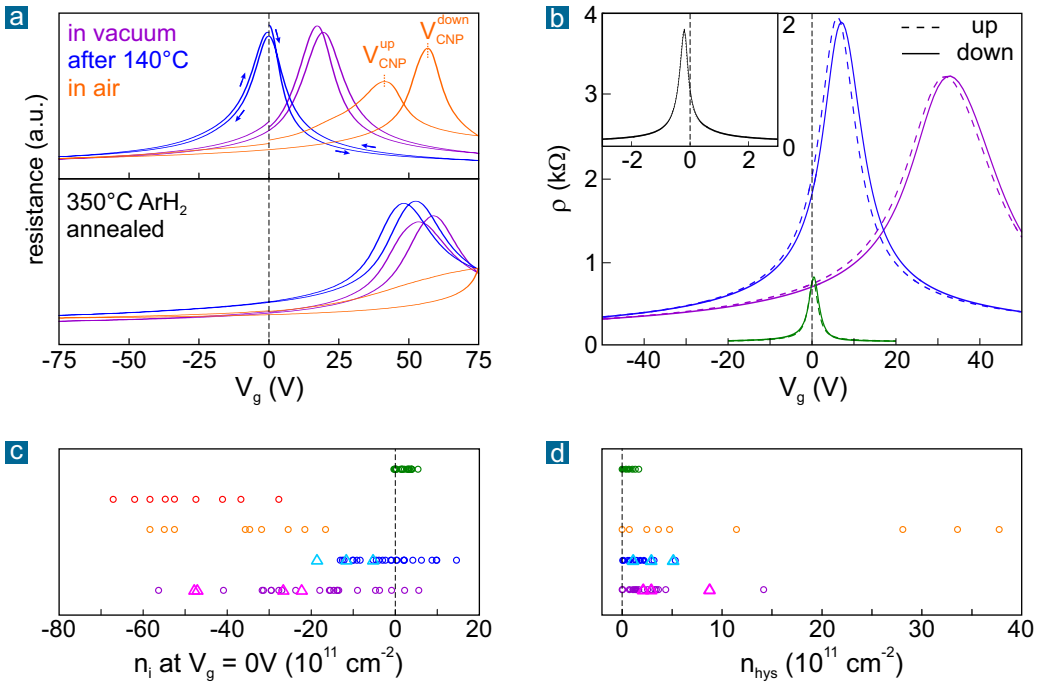

Figure 3.7. Field effect gating of graphene devices at room temperature and statistics on initial doping level and hysteresis. a: Longitudinal resistance of uncovered graphene on SiO_2/Si measured while sweeping the back gate voltage V_g upwards and downwards (starting and ending at zero gate voltage). The top panel shows an exemplary device without 350°C ArH_2 treatment, while the device in the bottom panel was annealed prior to the electrical measurements. The latter exhibits much stronger hole doping which persists even after baking the sample at 140°C in vacuum. The SiO_2 thickness is 290 nm and the gate factor determined by Hall measurements is $\beta = 7.4 \cdot 10^{10}\,\mathrm{cm}^{-2}/\mathrm{V}$. b: Resistivity curves for graphene on SiO_2/Si ($\beta = 7.2$) in vacuum (purple) and in vacuum after 140°C baking (blue), compared to a hBN-encapsulated graphene device (green) which is gated by the silicon back gate (290 nm and 40 nm hBN, $\beta = 5.86$). The inset shows another hBN-encapsulated graphene device tuned by a thin graphitic back gate (thickness of bottom hBN 13 nm, $\beta = 110$). c: Large ensemble statistics on the initial doping level n_i for different types of graphene devices (each data point corresponds to an individual sample). We distinguish hBN-encapsulated graphene (green, with or without a thin graphitic back gate) and uncovered, non-annealed graphene on SiO_2 in different environments (following the same coloring as in panel a). The red data points correspond to ArH_2 annealed graphene on SiO_2. Devices made from CVD grown graphene transferred to SiO_2/Si substrates are marked by triangles. d: Statistics on the hysteresis effect for the same devices as in panel c (see main text). For most of the annealed devices, the hysteresis could not be determined without leaving the safe window of the gate dielectric (hence, no data is shown here).

should be negligible. Moreover, both have to be independent of the environment.

Again, the key is proper[27] encapsulation. Our hBN-encapsulated devices show a remarkable decline of the initial doping level and, especially, of the sample to sample variation (see Table 3.1). In contrast to uncovered graphene on SiO_2, we do not observe extrinsic hole doping but, if at all, electron doping. This supports again the picture that the dangling bonds terminating the SiO_2 surface play a crucial role in the hole doping mechanism, as discussed in the previous section on Raman spectroscopy. Our assessment is that the slight electron doping stems from trapped hydrocarbons at the interface between graphene and hBN (and likely also between hBN and SiO_2). We also see a trend that annealing at higher temperature after stacking and AFM ironing results in devices with virtually zero initial doping.

The hysteresis is reduced by about a factor of five for hBN-encapsulated devices compared to uncovered samples. It is also stable and independent of the environment (e.g. in air or vacuum) for such encapsulated devices. Therefore, possible explanations of the hysteresis involve charged defects at the oxide interface and/or polar molecules, e.g. water, trapped at the interfaces and adsorbed to the graphene surface (in case of uncovered devices), as proposed in the literature [93, 94, 321–323]. Based on our data (Figure 3.7d and Table 3.1), the strength of the hysteresis effect is substantially increased in (humid) air for bare devices on SiO_2 which indeed points to adsorbates. These adsorbates seem to be loosely bound and volatile, since already repeating the measurement in vacuum reduces the hysteresis significantly. Additional baking in vacuum has then only a smaller impact on the hysteretic phenomenon (while it removes notably the initial doping). All in all, an improved gating behavior is only observed after encapsulation with hBN flakes and the extent of the effect tends to decrease with the cleanness of the graphene-hBN interface (that means with the annealing temperature after stacking). The residual hysteresis then seems to be caused by trapped charges at the hBN-SiO_2 interface since the implementation of a thin graphitic back gate (instead of the silicon back gate) leads to nearly perfect reproducibility of the field effect curves, based on three samples we have measured as well as reports in the literature (see inset of Figure 3.7b and [321, 322]). In particular, the authors of Ref. 321 claim an enhancement of the hysteresis effect at elevated temperatures (above room temperature) when the gating is conducted via the hBN-SiO_2 dielectric stack, while the graphitic back gate shows no (increase of the) hysteresis.

27 As already mentioned earlier, encapsulation with oxides [93] or resists [242] improves the extrinsic doping level and hysteresis but limits the electronic properties of graphene. Although placing graphene on hBN without capping layer already reduces the initial doping level [161], the uncovered surface of graphene still makes it sensitive to processing steps and harsh environments. Only encapsulation prior to further processing guarantees that the system is stable, e.g. even against plasma treatments [10].

type	n_i (10^{11} cm^{-2})	n_{hys} (10^{11} cm^{-2})
on SiO_2 in vacuum	-18.8 ± 14.8	2.5 ± 2.8
after 140°C baking	-1.5 ± 7.7	1.7 ± 1.3
350°C annealed	-49.8 ± 12.7	-
hBN-encapsulated	2.1 ± 1.7	0.47 ± 0.47

Table 3.1. Mean value and standard deviation of initial doping and hysteresis for different types of graphene devices. Negative n_i values correspond to initial hole doping, positive to initial electron doping.

We conclude that neither the initial doping nor the hysteresis are inherent to graphene but originate from extrinsic sources. They constitute a major challenge for engineering applications of graphene on a large scale. The key is encapsulation with a material that exhibits a small density of charged impurities and a high affinity for graphene so that the self-cleansing effect leads to atomically clean interfaces. This material then also serves as the gate dielectric. At present, hexagonal boron nitride stands out as the ideal partner for graphene. The encapsulation stabilizes the device's characteristics and makes them largely independent of the environment. Whether operated in air or in vacuum, they remain essentially the same. In the following chapters, if not specified otherwise, graphene on SiO_2 always refers to non-annealed but vacuum baked samples, and hBN-encapsulated devices use the integrated back gate of the silicon substrate (that means the dielectric stack consists of 290 nm SiO_2 and the bottom hBN flake). All measurements are conducted in vacuum. For other analysis, e.g Hall coefficient and noise measurements presented in the following chapters, we consider only one direction of the gate voltage sweep. We have arbitrarily chosen the upward sweep direction, starting from the most negative gate voltage.

Electrical transport at cryogenic and room temperature - effect of temperature on the residual density and mobility

As described in the introduction of this section, we can convert the measured resistivity field effect curves, e.g. shown in Figure 3.7b, into a double-logarithmic diagram of the conductivity over the gate induced density range (see Figure 3.8a). This allows us to deduce the residual total charge carrier density at the CNP and the mobility which constitute the main figures of merit. We present here two examples of a bare graphene on SiO_2 device (solid lines) and of a high quality, hBN-encapsulated device which was annealed at 500°C and AFM ironed (dashed lines). We further distinguish between room temperature and low temperature (10 K).

The remarkable difference between the two sample types is that the curves of the device on SiO_2 are almost identical at both temperatures[28]. This implies that the mobility and the residual carrier density hardly depend on temperature. The mobility is also density independent (which results in the observed $\sigma \propto n_g$ behavior over a wide density range up to high densities). This indicates that long range disorder, for instance generated by random strain fluctuations or charged impurities, represents the dominant scattering mechanism for the transport mobility (exceeding the phonon scattering contributions even at 300 K) and is also mainly responsible for charge carrier generation in the form of electron hole puddles (disorder induced charge carriers exceed the thermally activated ones).

In contrast, the conductivity of the hBN-encapsulated device alters considerably with temperature. The residual carrier density at the CNP decreases with temperature which points to a substantial contribution of thermal activation of charge carriers at room temperature (see also Table 3.2). The drop of the residual carrier density is also reflected in the reduction of the minimum conductivity (or the increase of the maximum resistivity) at the CNP when the temperature is lowered. In general, the conductivity of the hBN-encapsulated device is higher than for the device on SiO_2, as a consequence of the overall higher mobility. When the density is increased, we observe at both temperatures a change from a linear relation between conductivity and density to a sublinear trend.

At room temperature this behavior can be explained within the framework of the semiclassical model for diffusive transport (see Section 2.3.1). Due to the high sample quality, the long range disorder is so weak at higher densities that short range scattering finally dominates which leads to a density-independent conductivity (see Equation 2.51 and Figure 2.6). At room temperature and at high density, e.g. $n_g = 10^{12}\,cm^{-2}$, scattering at intrinsic acoustic phonons is known to be the ultimate limit for charge transport (see Section 2.3.2). The resistivity is around 50 Ω (and the conductivity is 20 mS, indicated by the violet line in Figure 3.8a)[29]. At

28 The wiggles seen in the low temperature curve at high density likely stem from conductance fluctuations. Note, the graphene on SiO_2 device presented here serves as a typical example representing the average device quality achievable on SiO_2. Some devices on SiO_2 do exhibit a (slight) temperature dependence of the mobility and residual density when the disorder induced electron hole puddle density is smaller (due to non-reproducible, fortuitous circumstances during device processing, as discussed in the first section of this chapter).

29 Equation 2.55 gives a resistivity of about 40 Ω assuming the acoustic deformation potential D_A equals 20.5 eV. However, in the literature a range of values are reported for this quantity (see Section 2.3.2). A revised theory even predicts the resistivity to be underestimated by ≈ 30%. This leads to 50 Ω [78]. This value is reached only for our very best sample and thus serves as a good estimate for the ultimate limit. The mobility is then calculated using Equation 2.38 and the mean free path is deduced from Equation 2.49 which is valid at high density ($k_BT \ll E_F$).

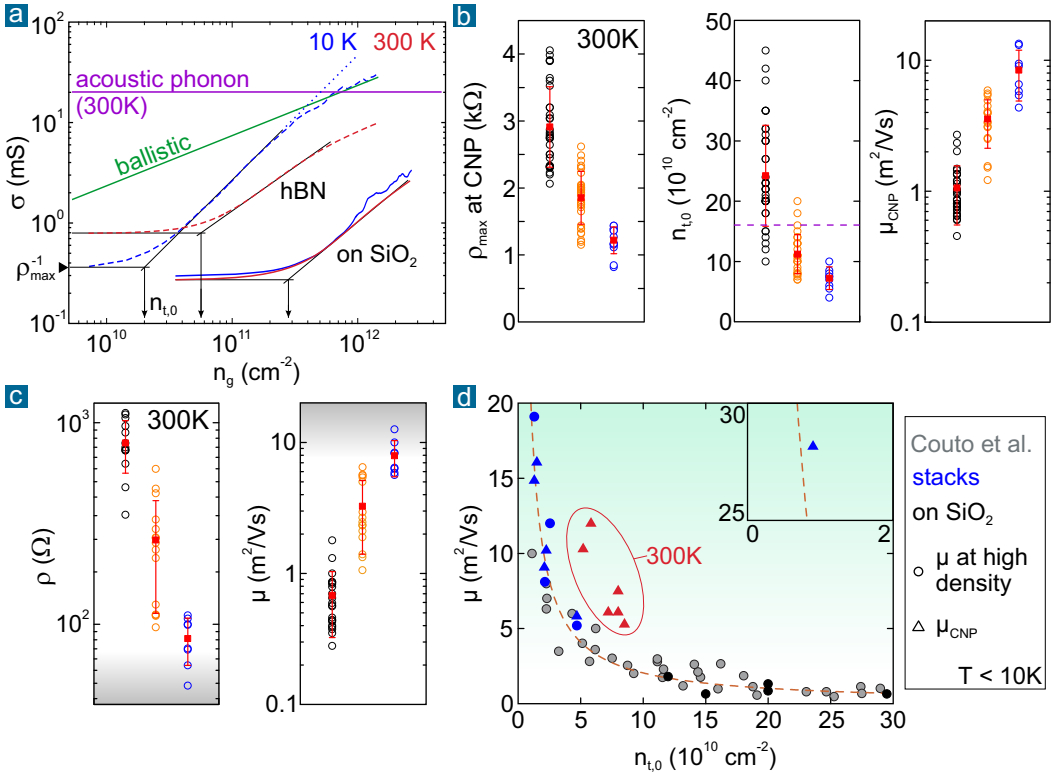

Figure 3.8. Statistics of the transport properties. a: Double-logarithmic diagram of the conductivity versus gate controlled average density. The residual total density at the CNP is determined as illustrated. b: Statistics of the room temperature resistivity maximum ρ_{max} and the residual total density $n_{t,0}$ at the CNP. We distinguish between graphene on SiO_2 (black) and hBN-encapsulated devices (orange: annealed at $\leq 350°C$, blue: at 500°C followed by AFM treatment). The violet dashed line marks the lower boundary due to thermal generation (for $v_F = 10^6$ m/s in Eq. 2.24). The last panel shows the mobility at the CNP (see main text). c: Room temperature resistivity at $n_g = 10^{12}\,cm^{-2}$ and the corresponding mobility (Eq. 2.38). Gray shaded areas mark the acoustic phonon scattering limit (Eq. 2.55). d: Correlation between mobility at higher density and residual density at the CNP at low temperature (circular data points). It indicates long-range disorder as the common origin for mobility reduction and residual carrier density. Data obtained on our hBN-encapsulated devices (blue) and devices on SiO_2 (black) are plotted together with the data of Couto et al. (gray) [18]. The dashed orange line represents the expected relation between $n_{t,0}$ and μ based on random strain fluctuations as dominant long range disorder (see Eq. 2.54). Deducing mobility values from the measured resistivity at high densities can become inaccurate in the shaded turquoise range owing to ballistic transport. Hence, we also calculated μ_{CNP} (triangles) for our best devices (see Table 3.2). At low temperature we prefer to determine $n_{t,0}$ from the Hall coefficient (see main text and Chapter 4).

$n_g = 10^{12}\,cm^{-2}$ (in the single carrier type regime), this corresponds to a mobility of about $125,000\,cm^2/Vs$ and to a mean free path L_{mfp} on the order of one micrometer. Considering our typical device dimensions of a few micrometer, our system is still in the diffusive regime. As the temperature is lowered, the contribution to the resistivity from acoustic phonon scattering decreases linearly with temperature (Equation 2.55). Hence, the mobility increases inverse with temperature. At low temperature the phonon contribution vanishes and we would expect the linear relationship between conductivity and density to hold at high densities since scattering is caused only by the weak long range disorder potential (dotted blue line in Figure 3.8a). However, our measurements clearly reveal a sublinear trend also at 10 K. In contrast to the room temperature case, it indicates a crossover from the inevitably diffusive regime around the CNP to the ballistic regime at high density[30]. To support this assertion, we have calculated the ballistic conductivity limit due to boundary scattering using Equation 2.71 and the channel width of the measured device (green line in Figure 3.8a). The width has been determined by AFM measurements. The agreement of the ballistic limit with the measured low temperature conductivity at high density is compelling.

Hence, one has to take into account a possible crossover to ballistic transport at low temperature when working with small sized devices. For example, Couto and coworkers claim that, in clean devices, random strain fluctuations dominate the long range disorder potential. The latter both imposes the electron hole puddle landscape (the residual density) and limits transport at low temperature (see Section 2.3.2 and Ref. 18). They determine the residual density at the CNP in the double-logarithmic plot of conductivity versus density which we have described above. The mobility is then calculated from the measured conductivity "at few $10^{11}\,cm^{-2}$". Their data points corresponding to non-encapsulated graphene on hBN or on SiO_2 devices (with few micrometer dimensions) are replotted as gray circles in Figure 3.8d. The theoretical model which describes the relation between residual density and mobility is plotted as orange dashed curve (see Equation 2.54 and note that n_* in the equation is equal to $n_{t,0}$ in the low temperature limit). We have included the results obtained on our devices in the same graph (devices on SiO_2 marked with black symbols and best hBN-encapsulated devices with blue symbols). Our best hBN-encapsulated devices, which were annealed at 500°C and AFM treated, are dominated by boundary scattering in the ballistic regime at high densities, e.g. $n_g \approx 10^{12}\,cm^{-2}$. Hence, the

30 The sublinearity in the high density regime at low temperature could also be explained by some "residual" short range disorder of unknown origin, within the diffusive picture (Sections 2.3.1 and 2.3.2). However, our explanation in terms of boundary scattering in the ballistic regime is compelling since independent measurements (of the bend resistance) demonstrate mean free path lengths of several tens of micrometer for hBN-encapsulated devices at low temperature [10, 195, 203]. These values clearly exceed our device dimensions.

mobility definition in the diffusive picture becomes meaningless and the calculated values may be inaccurate[31]. We therefore do not derive the mobility at higher densities but determine the (average) mobility at the CNP. The latter can be deduced from quantities which are obtained directly from the measured conductivity curves within the density range close to the CNP where we can be certain to be in the diffusive regime (as discussed thoroughly in Section 2.5). The minimum conductivity (the maximum resistivity) is directly measured and the residual density can be determined from a double-logarithmic diagram of conductivity versus density (or by fitting the two carrier type Drude model within channel approximation to the Hall coefficient curve[32]). The minimum conductivity at the CNP is then given by (see Equation 3.1)

$$\sigma_{\mathrm{min}} = \frac{1}{\rho_{\mathrm{max}}} = \mathrm{e}\,\mu_{\mathrm{CNP}}\,\mathrm{n}_{\mathrm{t},0}\,. \tag{3.4}$$

Since we have assumed electron and hole mobilities to be equal, the (average) mobility at the CNP μ_{CNP} was introduced. Note, that using the simple two carrier type Drude model also presumes the system to be spatially homogeneous (as when determining the residual density in the channel approximation introduced in Section 2.3.3). From a practical point of view, both quantities ($\mathrm{n}_{\mathrm{t},0}$ and μ_{CNP}) constitute excellent parameters to characterize the electrical transport through the electron hole puddle landscape. At low temperature, both parameters are governed by the long range disorder potential. The correlation between them yields valuable information about the dominant disorder mechanism. Our data points determined on extremely clean hBN-encapsulated devices (blue triangles in Figure 3.8d) extend the range of the non-encapsulated graphene devices of lower quality measured by Couto and coworkers in Ref.18. The proposed model involving random strain fluctuations shows good agreement with our measurements.

While the mobility and the residual total carrier density at the CNP correlate and are governed by only one common long range disorder mechanism at low temperature, the situation becomes more complex at room temperature. Due to graphene's

31 Ignoring this fact, we have converted the measured conductivity values into mobilities within the diffusive picture (blue circles). Surprisingly, those data points do not deviate substantially from the expected curve. This is however due to a fortunate coincidence. The ballistic conductivity at $10^{12}\,\mathrm{cm}^{-2}$ for our typical sample width is similar to the conductivity which we would expect if the long range scattering potential would matter (dotted blue line in Figure 3.8a). For example, for $\mathrm{W} = 2\mu\mathrm{m}$ the ballistic conductivity allows us to derive a mobility of $109,000\,\mathrm{cm}^2/\mathrm{Vs}$ (in the incorrect diffusive picture) which would also fit to the theory.

32 The fitting procedure is described in detail in Chapter 4. Both methods often lead to only small deviations in the derived $\mathrm{n}_{\mathrm{t},0}$ values (for the same sample). However, in particular at low temperatures, we prefer the automated fitting of the Hall coefficient over the semi-manual determination of the $\mathrm{n}_{\mathrm{t},0}$ value in the double-logarithmic method (which may reduce the precision due to reading errors).

device	$n_{t,0}$(10K)	μ_{CNP}(10K)	$n_{t,0}$(300K)	μ_{CNP}(300K)
S1	0.9*	28*	5.82	12.0
S2	1.3	14.9	5.2	10.3
S3	1.5	16.1	8.0	7.5
S4	2.1	9.1	7.2	6.1
S5	2.2	10.2	8.0	6.1
S6	4.7	5.8	8.5	5.3

Table 3.2. Residual total charge carrier density (in $10^{10}\,cm^{-2}$) and mobility at the CNP (in m^2/Vs) for the best hBN-encapsulated devices (subject to the post-assembly treatment). Here, $n_{t,0}$ is derived from the Hall coefficient curve (as discussed in detail in Chapter 4). The latter allows an automated fitting with a better precision than the (semi-manual) double-logarithmic method. Both methods, on average, exhibit only minor deviations in the $n_{t,0}$ values (*= determined by double-logarithmic method only).

gapless band structure, charge carriers are thermally activated which leads to an increase of the residual density at the CNP (see Section 2.2.2)[33]. For our cleanest devices presented in Table 3.2, we observe an up to 5-fold increase of the residual density at room temperature compared to low temperature. This indicates that the thermal excitation flattens the electron hole puddle landscape making the system more homogeneous. While the gain in the total carrier density at the CNP mostly causes the rise of the minimum conductivity (the drop of the maximum resistivity), we find the average mobility at the CNP to decline with increasing temperature (and to decline comparatively stronger for the very best devices with the highest low temperature mobility, as listed in Table 3.2). Hence, for clean devices, the correlation between $n_{t,0}$ and μ_{CNP} at room temperature cannot be used to test the common long range disorder mechanism theory (see red triangles in Figure 3.8d).

The temperature dependence of the mobility can be explained by scattering off of graphene's intrinsic phonons and/or by the temperature dependence of the long range disorder scattering: While acoustic phonon scattering dominates at high density (in the single carrier type regime), the optical phonon mode A_1' has been identified recently as the major contribution to phonon scattering at low carrier density, e.g. in vicinity of the CNP (Section 2.3.2 and Ref. 78, 80). Moreover, as already discussed in Section 2.3.2, scattering caused by the long range disorder potential

33 In case of the linear band structure with Fermi velocity $v_F = 10^6\,m/s$, we would expect a residual density of $16 \cdot 10^{10}\,cm^{-2}$ at room temperature (Equation 2.24 and violet line in Figure 3.8b). Surprisingly, we measure a substantially lower residual density for most (clean) devices, as presented in Table 3.2 and Figure 3.8b. We will address this issue in more detail in the following chapter. In our present discussion on sample statistics we simply continue to work with the measured values.

is weakly temperature dependent as a consequence of the temperature dependence of graphene's dielectric screening function [80, 102, 109, 180]. Whether the mobility drops substantially with increasing temperature, in case of dominating optical phonon scattering, or just weakly, in case of dominating disorder scattering, depends on the disorder level of the studied sample. The lower the disorder level, the lower is the temperature threshold above which optical phonon scattering prevails. Sample S1 with the lowest disorder level (that means the highest mobility at low temperature) exhibits the largest drop of the mobility at the CNP when increasing the temperature from 10 K to 300 K. This is because the effect of phonon scattering is larger. In contrast, sample S6 with a higher disorder level exhibits only a minor decrease in mobility since long range scattering remains dominant even at room temperature.

In conclusion, the temperature dependence of the minimum conductivity at the CNP arises from the competition of both effects: an increase of the residual density at the CNP and a reduction of the mobility with temperature. The stronger the extrinsic long range disorder, e.g. for devices on SiO_2, the weaker the change of the conductivity at the CNP with temperature. Both the residual density at the CNP and the mobility are almost independent of temperature and cancel each other within the typical experimental temperature range up to room temperature (or slightly above). Hence, as we will discuss in detail in Chapter 4, we observe only for our hBN-encapsulated devices with lowest disorder level a pronounced insulating behavior at the CNP. It can be mainly attributed to the increase of the residual carrier density at the CNP due to thermal activation of carriers. The transition from insulating behavior ($\partial\rho_{max}/\partial T < 0$) to metallic ($\partial\rho_{max}/\partial T > 0$) behavior finally occurs above room temperature. Then, the mobility decrease due to phonon scattering ultimately exceeds the effect of thermal generation of additional carriers.

Statistics on residual density and mobility at room temperature - evidence for the beneficial effect of the proposed post-assembly treatment

We now attempt to answer the questions, whether the intrinsic performance limits at room temperature can be reached reproducibly by following a certain processing procedure, what these limits actually are, and how we can easily judge the quality of a device through fast electrical measurements, such as for instance the measurement of the resistivity values at the CNP and at high density ($n_g = 10^{12}\,cm^{-2}$) at room temperature. We have compiled these two resistivity values and the residual density

at the CNP of many devices in Figure 3.8 b and c, respectively[34]. The dimensions of all devices are large enough such that the picture of diffusive transport holds over the whole density range (at room temperature). This allows us to define and deduce mobility values at the CNP (Figure 3.8b) and at high density in the single carrier type regime (Figure 3.8c). As expected from our pre-assessment using Raman spectroscopy, bare graphene devices on SiO_2 perform poorly and all parameters are governed by extrinsic disorder and Coulomb impurities constitute the main contribution to the strong long range disorder potential. This is reflected in the low mobility of about $10,000\,\mathrm{cm}^2/\mathrm{Vs}$ and the large residual density of $n_{t,0} = 24 \cdot 10^{10}\,\mathrm{cm}^{-2}$. Also not surprisingly, the device quality improves after encapsulation of graphene in between hexagonal boron nitride flakes.

More intriguing, however, is our finding that hBN-encapsulated devices can be substantially and reliably improved by conducting the additional post-assembly treatment described in Section 3.1.5 (see also our work published in Ref. 183). In general, we see a clear tendency that a higher annealing temperature after stack assembly reduces the residual density and increases the mobility. This points to the importance of the interface cleanness on the atomic scale (beyond the resolution limit of atomic force microscopy which we use to avoid bubbles). Trapped contaminants, likely a major source of charged impurity scattering and interface roughness, are efficiently removed by thermal facilitation of the self-cleansing effect. The next leap forward in the device performance is then unambiguously achieved by applying the AFM ironing technique after all other device processing steps have been completed. As discussed already in Section 3.1.5, we surmise that the underlying mechanism is smoothing nanometer scale corrugations at the interfaces. This reduces random strain fluctuations which constitute the remaining predominant source of disorder when the contribution of charged impurities is already minimized. Hence, the effectiveness of the AFM treatment is found to rely on stacks which already possess an excellent interface cleanness. That means that we typically observe a quality enhancement due to the AFM treatment only for devices that have been annealed at 500°C after stack assembly (see Table 3.3 and Figure 3.9). In conclusion, the high-temperature annealing is an important pre-requisite and a reliable improvement requires combining both steps.

34 The majority of our graphene on SiO_2 devices are measured in a two terminal (Hall cross) configuration. However, the contribution of the contact resistance is minor and causes only a small error when determining the resistivity (and from that the mobility), especially at the CNP. We also measured several four-terminal devices which show similar values and spread as seen for the two-terminal geometries. Only at high density, the resistivity is, on average, slightly lower (e.g. about $500\,\Omega$ instead of $800\,\Omega$). This results in a marginally higher mobility of about $10,000\,\mathrm{cm}^2/\mathrm{Vs}$ (instead of $6800\,\mathrm{cm}^2/\mathrm{Vs}$) which, however, does not alter any conclusions we make here.

°C	μ_{FE} before	$n_{t,0}$ before	μ_{FE} after	$n_{t,0}$ after
200	2.5 (2.3)	12	2.5 (2.3)	12
350	6.4 (5.7)	9.1	6.4 (5.6)	8.5
500	6.7 (6.3)	11.3	10.1 (8.3)	6.3

Table 3.3. Field effect mobility (as defined in Equation 2.38) and residual carrier density at the CNP *at room temperature* before and after AFM treatment conducted as final step. We distinguish between three devices which have been annealed at different temperatures (first column) after stacking. Mobilities are determined at $n_g = 5 \cdot 10^{11}\,cm^{-2}$ ($1 \cdot 10^{12}\,cm^{-2}$) and given in units of m^2/Vs. The residual carrier densities are given in $10^{10}\,cm^{-2}$ and determined by fitting the Hall coefficient curves (see next chapter). The corresponding electrical measurements of the Hallbar shaped devices are shown in Figure 3.9.

Accordingly, we have grouped stacks which were solely annealed at 200°C or 350°C (orange points) as well as stacks which were annealed at 500°C and then AFM treated after device processing (blue points) in our statistics presented in Figure 3.8. Only in the latter case we reproducibly obtain mobility values at high density which approach the limits set by intrinsic acoustic phonon scattering. On average, the measured resistivity at room temperature and at $n_g = 10^{12}\,cm^{-2}$ is $85\,\Omega$ which converts to a mean mobility of about $80,000\,cm^2/Vs$ ($8\,m^2/Vs$). These optimized devices then typically possess a maximum resistivity around $1000\,\Omega$ and a residual carrier density around $n_{t,0} = 7 \cdot 10^{10}\,cm^{-2}$ at the CNP. Since we know from the former discussion that the residual density drops considerably when the temperature is lowered, the residual carrier density at room temperature is determined by thermal activation of charge carriers and thus constitutes an intrinsic limit (see also Section 4.3). The estimated mobility at the CNP, based on our definition in Equation 3.4, is almost identical to the high density value ($\bar{\mu}_{CNP} = 84,000\,cm^2/Vs$). The match is probably coincidental since, in theory, the mobility at the CNP is expected to be limited by different scattering mechanisms (long range disorder and intrinsic optical phonons) than the mobility at high density (intrinsic acoustic phonons). In conclusion, devices with excellent electrical properties can be identified by overall lower resistivity values at room temperature[35]. A fast and, as we have shown, statistically reliable method to assess the sample quality is to measure the resistivity at the CNP and at high density. Room temperature performance limits are reached when the maximum resistivity is found to be near or even below $1000\,\Omega$ and the resistivity in the high density limit to be below $100\,\Omega$ (and best values come close to $50\,\Omega$, the intrinsic limit set by acoustic phonon scattering). As a rule of thumb, the ratio between the two resistivity values should be 10:1. The here and in our publication

35 We emphasize that the maximum resistivity of high quality samples is lower despite the smaller residual density. This is due to the strong gain in mobility for clean devices.

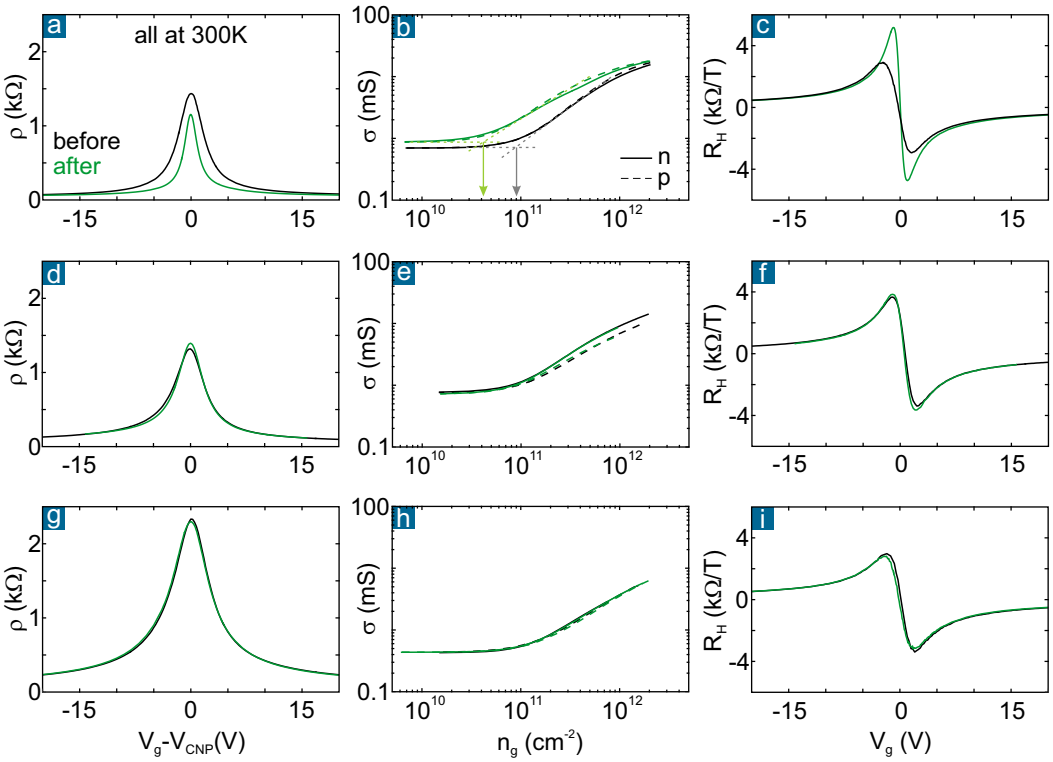

Figure 3.9. Effect of the final AFM ironing step performed on three hBN-graphene-hBN stacks which were annealed at different temperatures subsequent to their assembly. The device presented in the top row (a-c) was annealed at 500°C, the device in the middle (d-f) at 350°C, and the device in the bottom row (g-i) at 200°C. Black (green) curves are measured at room temperature before (after) conducting the AFM treatment. Resistivity over back gate voltage (initial doping is corrected) is shown in the first column. These curves are converted into a double-logarithmic diagram of the conductivity versus the gate controlled average density (second column). This allows us to make a reasonable estimate of the residual total charge carrier density at the CNP before and after AFM treatment (arrows). The last column presents the Hall coefficient versus the effective back gate voltage. The shape of the curve and the Hall coefficient extrema change considerably after applying the AFM ironing only for devices annealed at 500°C.

(Ref. 183) proposed post-assembly treatment increases substantially the yield of devices that approach these intrinsic limits at room temperature.

Chapter 4

Weak field magneto-transport

In the previous chapter we have demonstrated that electrical transport measurements in the absence of a magnetic field already allow us to (pre-)characterize graphene devices, e.g. in terms of mobility and residual density at the charge neutrality point. However, in order to determine these two quantities from resistance measurements, we rely on knowing the exact carrier density at a certain applied gate voltage. The gate (conversion) factor β can either be calculated when the gate geometry and dielectric material are well known, or it is determined unambiguously through Hall measurements. Vice versa, if the gate factor is known, graphene constitutes a system whose carrier density can be tuned precisely through applying a gate voltage. Comparatively small density values, that means large Hall coefficient values, can be reached while the system remains well conducting (due to the high mobility even at room temperature). This combination makes graphene a promising material for magnetic field sensing applications which we will elaborate further from a technical point of view in the next chapter. In this chapter we report on our magneto-transport measurements in classically weak magnetic fields, i.e. when $\mu^2 B^2 \ll 1$ holds. For the typical mobility μ of our samples this implies a magnetic field B_{max} of about 50 mT. In particular, we examine the behavior at room temperature as relevant for the envisaged use in applications. Hence, we exclude here quantization effects and quantum interference effects. They are discussed briefly in appendix B.1. Although, at first glance, this study may seem to deal with rather trivial physics, graphene turns out to be a challenging system to understand thoroughly when we dig deeper into the details which otherwise are typically rushed through in the literature. At the end, however, this understanding might be decisive for graphene's future in applications involving markets with billions worth of revenue as targeted by the EU graphene flagship program.

Graphene provides a beautiful textbook example of a gate tunable crossover from a single carrier type system (at high density) to a two carrier type system (close to

the CNP). Aside from the finite residual conductivity, the two carrier picture is supported by observing a zero Hall coefficient and strong classical magneto-resistance at the charge neutrality point. We will demonstrate that both phenomena can be (at least qualitatively) explained by the simple two carrier type Drude model which assumes the system to be spatially homogeneous (as in the case of a compensated semiconductor, see Section 2.3.3). How well this condition is fulfilled depends on the strength of the extrinsic disorder, which imposes a spatially inhomogeneous landscape of electron and hole puddles, and on the actual temperature (or, to be more exact, on the ratio between temperature and the disorder's effective temperature scale). The band structure of disorder and strain free graphene - that is under ideal conditions from a theoretical point of view - is symmetric and exhibits no band gap. Hence, electrons and holes are generated equally by thermal excitation. The scattering mechanisms are not sensitive to the carrier type which implies equal electron and hole mobilities. All in all, this makes graphene the most trivial example of a compensated semiconductor, being spatially homogeneous and, thus, completely captured by the two carrier type Drude model. However, this ideal picture is approached in reality only for devices of excellent quality at elevated temperatures, e.g. for hBN-encapsulated graphene. Since the extrinsic disorder level is outstandingly low, thermal generation then already dominates at room temperature and the system can be considered as quasi-homogeneous. In case of low sample quality (stronger extrinsic disorder) and/or in the low temperature limit, the system eventually becomes spatially inhomogeneous. The two carrier type Drude model then only serves as an approximation which, however, is found to work quite well when one introduces the residual total carrier density at the CNP as if the system were homogeneous. This can be justified by the fact that the boundaries between the electron hole puddles are highly conducting and quasi transparent.

The chapter is split into six sections. The first three sections discuss the most important results regarding graphene's employment as transducer material in Hall sensors. The first section is dedicated to the Hall effect in graphene at room temperature and explains how we obtain the Hall coefficients as a function of gate voltage. We apply the simplest form of the two carrier type Drude model using the channel approximation in order to fit the gate voltage dependence of the Hall coefficient, which we refer to as the "Hall coefficient curve" in the following. This approach allows us to deduce the residual carrier density at the CNP in an independent manner (as an alternative to the semi-manual analysis in the double-logarithmic diagrams of the conductivity versus density). Our study also encompasses a device statistics which compares different device types, e.g. non-encapsulated and encapsulated samples, with respect to the residual density at the CNP and the maximum Hall coefficient (which is equivalent to the maximum sensitivity, a key parameter for Hall sensors). Furthermore, we compare the applicability of the two carrier type Drude model's simplest form

with a minimal number of fitting parameters for lower quality (on SiO_2) and for higher quality (hBN-encapsulated) samples. The minor but yet noticeable asymmetries we see between the electron- and hole-side in some of our devices, in particular for the Hall coefficients at lower temperature and in case of lower quality samples with stronger disorder level, are covered in the second section. While asymmetries close to the charge neutrality point can be explained by unequal carrier mobilities, the origin of the observed asymmetries at high density in the single carrier type regime remains obscure. We debate the introduction of individual Hall scattering factors for the electron and hole side in order to be able to adapt the two carrier type Drude model. We execute different fitting routines which allow us to quantify error bars for the obtained residual carrier density values at the CNP. Surprisingly, we measure larger Hall coefficients and obtain smaller residual densities at the CNP than we would expect when simply calculating the thermal generation of carriers based on the linear band structure with a Fermi velocity of 10^6 m/s. This deviation is examined in more detail in the third section where we present the temperature dependence of magneto-transport in graphene (for weak magnetic fields). In particular, we investigate the temperature dependence of the maximum Hall coefficient and of the maximum resistivity from which one can deduce the temperature dependence of the residual density and the mobility at the CNP. In order to complete our picture, we discuss also the temperature dependence of the resistivity in the high density regime which is governed solely by the temperature dependence of the scattering mechanism, e.g. phonon scattering.

In the fourth section we discuss Hall measurements on bilayer and multilayer devices which exhibit much smaller maximum Hall coefficients at room temperature due to stronger thermal excitation of carriers which reflects their parabolic band dispersion. Furthermore, the effect of a Moiré superlattice on the transport properties is examined. The Moiré pattern occurs when the crystallographic planes of graphene and hexagonal boron nitride are properly aligned. In particular, we analyze the temperature dependence of the features at the main CNP and at secondary neutrality points which stem from band structure reconstructions due to the zone folding induced by the weak periodic potential. The fifth section treats the strong classical magneto-resistance which appears at and close to the charge neutrality point and may offer another possibility to measure magnetic fields using graphene devices. While the strong positive magneto-resistance at the CNP is understood (qualitatively) within the two carrier type Drude model, we also observe a negative magneto-resistance at higher densities when the temperature is lowered. This phenomenon is accompanied by a sublinearity in the Hall resistivity. In the sixth section we explore the potential origin of these findings by comparing our measurements to the theory in the ballistic and hydrodynamic transport regimes.

4.1. Hall effect and gate tunable Hall coefficient

4.1.1. Measuring principle and determination of the Hall coefficient

According to the semi-classical theory[1], in the single carrier regime at higher densities, we expect to observe a Hall voltage between transverse voltage probes but no longitudinal magneto-resistance effect when applying a magnetic field perpendicular to graphene's plane. The Hall voltage U_H is proportional to the current I and the magnetic field B, and the proportionality constant, the Hall coefficient R_H (in units of Ω/T or V/AT), scales inversely with the (average) carrier density (Eq. 2.40). We obtain the Hall coefficient from a linear fit[2] of the measured transverse (or Hall) resistivity, $\rho_{xy} = U_H/I$, as a function of magnetic field, as demonstrated in Figure 4.1a. The Hall coefficient, i.e. the slope, is positive (negative) for holes (electrons). At higher densities in the single carrier regime, the linearity of the measured data points is excellent over the whole magnetic field sweep range, in our case on the order of $\pm 200\,\mathrm{mT}$. Near the CNP, when electrons and holes coexist, the Hall resistivity curve can become non-linear because the absolute value of the Hall coefficient increases with the magnetic field when $\mu^2 B^2 > 1$ (see further explanation at the end of this subsection). Moreover, at lower temperature, non-linearity might also arise due to mesoscopic transport phenomena (as discussed in Section 4.6). Hence, in order to be sure that the extracted Hall coefficient value $R_H(\mu^2 B^2 \ll 1)$ is not affected by other physical phenomena for the hBN-encapsulated devices with high mobility, we restrict us to a smaller magnetic field range ($\pm 50\,\mathrm{mT}$) when analyzing the Hall resistivity curve $\rho_{xy}(B)$.

In our magneto-transport measurements, we tune the average carrier density n_g by changing stepwise the gate voltage V_g. For each voltage step, we sweep the

1 See Section 2.3.1 and Appendix B.2 with Figure B.2 sketching our typical Hall cross or Hall bar device geometry.

2 From the linear fit we also obtain the Hall resistivity offset ρ_{off} which we define as the transverse resistivity at zero magnetic field. If the zero magnetic field point is well known, e.g. through a reference magnetic field sensor, the offset can be explained by a slight misalignment between the transverse contacts, such that a parasitic, longitudinal voltage drop is measured in addition to the Hall voltage [150]. Moreover, it can be caused by inhomogeneities in the sample [347]. In general, we find the Hall offset to be larger near the CNP. It is an important parameter when applying graphene as Hall sensor transducer (see Chapter 5). However, it is not important for the general physics discussed here and the determination of the Hall coefficient. Hence, we subtracted for each gate voltage increment the individual offset from the respective Hall resistivity curve, resulting in Figure 4.1a. In doing so, the crossover from positive to negative slope as well as the almost horizontal curve close to the CNP ($R_H \to 0$) is brought out more nicely.

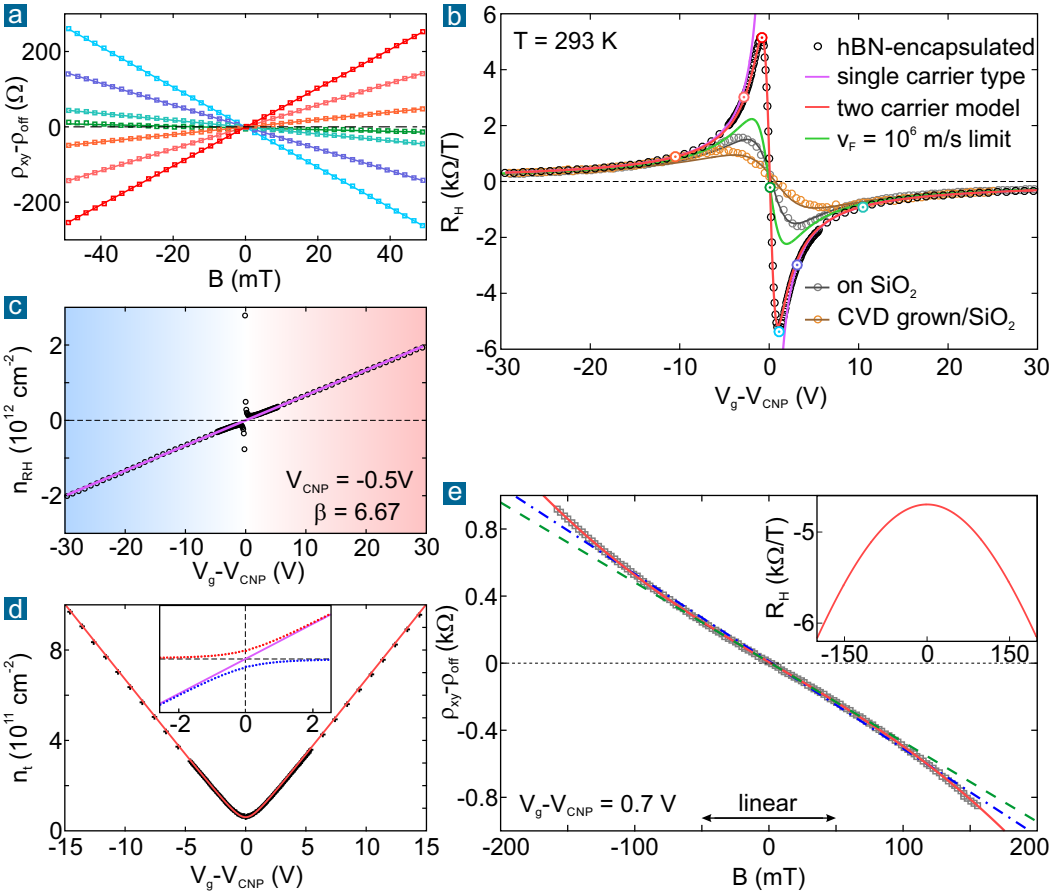

Figure 4.1. Hall effect in graphene at room temperature. a: Measured transverse resistivity (squares) during magnetic field sweeps for several gate voltages. Solid lines are linear fits to determine the Hall coefficients which are plotted in panel b against effective gate voltage (the Hall coefficients obtained from the Hall traces in panel a are marked by the highlighted circles of corresponding color). We distinguish hBN-encapsulated graphene (black), bare graphene on SiO_2 (gray) and CVD grown graphene transferred to SiO_2 (yellow circles). Fits using the two carrier type Drude model within the channel approximation are drawn as solid lines. The green line marks the expected thermal limit. c: Carrier densities derived from the Hall coefficients of hBN-encapsulated graphene using the single carrier formula (best fit violet line). d: Electron (red) and hole densities (blue points in the inset), and the respective total carrier density (black crosses) can be deduced from the determined Hall coefficient using the two carrier type Drude model (see main text). Y-axis range of inset is $\pm 2 \cdot 10^{11}\mathrm{cm}^{-2}$. The red solid line corresponds to the fit in panel b. e: Close to the CNP, deviations from the linear Hall traces arise for high mobility devices when the magnetic field strength is increased beyond the classical weak field limit ($\mu^2 B^2 \geq 1$). This may lead to an overestimation of the Hall coefficient when employing linear fitting (see main text).

magnetic field while recording the Hall voltage as well as the longitudinal voltage drop (from which we calculate the longitudinal resistivity at zero magnetic field)[3]. Panel b of Figure 4.1 shows the obtained Hall coefficients as a function of applied gate voltage, centered around V_{CNP}. In the following we will refer to $R_H(V_g)$ as the "Hall coefficient curve". In panel b, we present exemplary room temperature data obtained on a hBN-encapsulated graphene device, a bare graphene device on SiO_2, and a bare device fabricated from CVD grown graphene that was transferred onto SiO_2. The individually colored circles belonging to the Hall coefficient curve of the hBN-encapsulated device mark the Hall coefficient values that were obtained from the Hall resistivity traces plotted with the corresponding color in panel a of the figure. At higher gate voltages, that means higher densities, the Hall coefficients of all device types are equal and agree well with the single carrier type picture (Eq. 2.40 assuming for the Hall scattering factor $r_H = 1$, violet line in panel b). Vice versa, using the single carrier type equation with $r_H = 1$, the obtained Hall coefficients can be directly converted into density values ($n_{RH} = -1/e\,R_H$). Panel c of Figure 4.1 shows the calculated density values n_{RH} versus gate voltage, using the obtained Hall coefficients of the hBN-encapsulated graphene device. Taking into account only the high gate voltage regime on the hole (blue shaded) and electron (red shaded) side, we find $\beta = 6.67 \cdot 10^{10}\,\mathrm{cm}^{-2}/\mathrm{V}$ for the gate factor of this device[4]. The linear fit (violet line) is excellent both on the hole and electron side confirming electron-hole symmetry (see Section 4.2 for a discussion of potential asymmetries). The gate voltage corresponding to the CNP, V_{CNP}, is determined either from the maximum of the longitudinal resistivity or is used as an additional fitting parameter of the $n_{RH}(V_g)$ dependence. Finally, knowing both values allow us to convert precisely the applied gate voltage into the gate controlled average density n_g (Eq. 2.18).

Contrary to the prediction of the single carrier type model, all experimental Hall coefficient curves in Figure 4.1b eventually cross through zero at the CNP and are fairly point symmetric with respect to the latter. This provides evidence of the co-existence of electrons and holes near the CNP. As already discussed in detail in Section 2.3.3 and Appendix B.2, the two carrier type Drude model can be employed to describe to good approximation the magneto-transport in graphene, even if electrons and holes coexist in spatially allocated puddles. The two carrier type Drude model assumes a homogeneous system that is described by average electron and hole densities $n_{e,h}$ and mobilities $\mu_{e,h}$ (see Eq. B.16 for the general expression of the

3 While in the single carrier regime no magneto-resistance is observed in agreement with the semiclassical theory, a strong classical magneto-resistance arises near the CNP due to the coexistence of electrons and holes, as discussed in Section 4.5. The magnetic field dependence of the longitudinal resistivity is always measured in parallel to the Hall voltage. However, in this section, we consider only the longitudinal resistivity value at zero magnetic field.

4 If not otherwise specified, we will use $10^{10}\,\mathrm{cm}^{-2}/\mathrm{V}$ as the unit of the gate factor in the following.

Hall coefficient). Assuming equal mobilities and considering the weak magnetic field strength in our experiments, such that $\mu^2 B^2 \ll 1$, the Hall coefficient formula in the two carrier type Drude model breaks down to the simple Equation 2.62, which we repeat here for convenience:

$$R_H(\mu^2 B^2 \ll 1) = \frac{n_h - n_e}{e\,(n_h + n_e)^2} = -\frac{n_g}{e\,n_t^2}\,. \tag{4.1}$$

As aforementioned, the average carrier density $n_g = n_e - n_h$ is set by the applied gate voltage (see Eq. 2.18). Since we have already determined the gate factor from fitting the single carrier type model to the high gate voltage regime (see panel c), $n_g(V_g)$ is well known. Using the total carrier density $n_t = n_e + n_h$ and $n_g = n_e - n_h$, Eq. 4.1 can be rearranged in order to obtain direct expressions for $n_{e,h}(R_H, V_g)$:

$$n_h = 0.5\,(\sqrt{-n_g/eR_H} - n_g) \quad n_e = n_g + n_h\,. \tag{4.2}$$

The inset of panel d of Figure 4.1 shows the resulting electron (red dotted line) and hole (blue dotted line) densities considering the measured Hall coefficients of the hBN-encapsulated device. The average density, $n_g = n_e - n_h$, is given by the violet line. Obviously, it also represents the single carrier type model, where $n_t = n_g$. The main diagram of the panel then displays the directly derived total carrier density $n_t = n_e + n_h$ as a function of gate voltage (black crosses). The red solid line in panel d (and b) corresponds to the two carrier type Drude model fit using the channel approximation and the residual total carrier density at the CNP $n_{t,0}$ as the only fitting parameter. Due to its importance for this work, this approach will be discussed in detail in the next subsection 4.1.2. Asymmetries in the Hall coefficient curve between the electron and hole side can be properly addressed by this approach (see Section 4.2).

Exactly at the CNP, the coexistence of electrons and holes in equal numbers always leads to a zero Hall coefficient. Near the CNP, however, electrons and holes still coexist but in unequal quantities. The two carrier type Drude theory then predicts a magnetic field dependence of the Hall coefficient (see Eq. 2.61)[5]. This dependence becomes evident as soon as the condition $\mu^2 B^2 \ll 1$ no longer holds and causes the aforementioned non-linearity in the Hall resistivity versus magnetic field curves. For our hBN-encapsulated devices mobility values are typically above $50,000\,\mathrm{cm}^2/\mathrm{Vs}$ (even at room temperature). Hence, already magnetic fields on the order of 100 mT

5 We emphasize that the magnetic field dependence of the Hall coefficient is a direct consequence of the presence of electrons and holes. In the single carrier regime, the Hall coefficient is independent and the Hall response remains linear over the whole magnetic field range (except for mesoscopic transport effects at lower temperatures which we discuss in Section 4.6).

can no longer be considered as weak. In panel e of Figure 4.1 we illustrate this behavior for the same hBN-encapsulated device that we presented in the other panels of the figure. We have chosen a small effective gate voltage $V_g - V_{CNP}$ such that the system is close to the CNP but we still measure a Hall coefficient substantially different from zero (close to the maximum). The measured non-linear Hall resistivity trace (gray data points) is described nicely by the two carrier type Drude model (red solid line, Eq. 2.61) using the $n_{e,h}$ values that we obtained above for the effective gate voltage value. We assumed an equal mobility for electrons and holes and used it as the only fitting parameter. Its value is found to be comparable to the value we obtained from the field effect mobility equation. The inset of panel e depicts the magnetic field dependence of the Hall coefficient (obtained from the fitting procedure). We find $R_H(B \to 0) = -4697\Omega/T$. This value deviates only by about one percent from the Hall coefficient $R_{H,50mT} = -4755\Omega/T$ which we obtain by applying a linear fit to the data within a smaller magnetic field range of $\pm 50\,mT$ (green dashed line). However, when we ignore this issue and perform a linear fit over the whole magnetic field range of about $\pm 150\,mT$ (blue dashed line), the resulting Hall coefficient value is $R_{H,150mT} = -5203\Omega/T$. Not only the quality of the fit is much worse but also the value of the Hall coefficient is substantially overestimated. In particular, this may lead to erroneous statements on the maximum Hall coefficient. To conclude, our approach to limit the magnetic field range, guarantees to obtain the correct Hall coefficient value $R_H(\mu^2 B^2 \ll 1)$ over the whole gate voltage regime and for all device types (see device statistics with respect to the maximum Hall coefficient presented in Section 4.1.3).

Experimental details

Our magneto-transport measurements are performed either in a commercial, dry Oxford cryostat system, operating from 1.5 K to 300 K, or in a self-designed setup using an electromagnet and allowing a temperature range from room temperature (293 K) up to about 80 °C (370 K). The latter setup will be presented in detail in the next chapter where it will be used to perform electrical noise measurements. The electromagnet exhibits a very low magnetic field hysteresis and, in addition, we employ a reference magnetic field sensor. In doing so, the magnetic field value is well known which allows us to determine precisely the conductivity at zero magnetic field and the Hall resistivity offset. In the electromagnet setup, all magneto-transport measurements are conducted applying a direct current and multimeters for voltage measurements. Typically, a current on the order of $1\,\mu A$ is chosen. On the contrary, an AC configuration and lock-in amplifiers are employed in the cryostat system. Room temperature measurements performed on some selected devices

in both systems were found to coincide. This proves the reliability and correctness of our results. Most of the data sets at room temperature were acquired in our self-designed setup (e.g. the statistics in Section 4.1.3), while all measurements below room temperature were conducted in the cryostat system. The latter, however suffers from a small but for our purpose notable hysteresis effect when sweeping the magnetic field. This is related to trapped flux in the superconducting magnet. Moreover, a reference magnetic field sensor was not integrated into the system. We tackle this problem by considering always the same sweep direction for which we found the magnetic field offset to be sufficiently stable. Its absolute value then can be determined, for instance, by measuring the strong positive magneto-resistance at the CNP. It shows a clear, quadratic field dependence possessing a minimum at zero magnetic field. Accordingly, the conductivity at (the real) zero magnetic field and the Hall offset can be determined. Moreover, for the purpose of obtaining the Hall coefficient in the limit of classically weak magnetic fields, i.e. when the Hall resistivity curve is linear, the magnetic field hysteresis is insignificant.

4.1.2. Modeling of the Hall coefficient curve using the two carrier type Drude model and the channel approximation

In the two carrier type Drude model, the Hall coefficient for weak magnetic fields offers the advantage that it is determined solely by the electron and hole densities but does not depend on the absolute mobility values[6]. In contrast, the longitudinal resistivity at zero magnetic field is always a function of the carrier densities and mobilities (Eq. 2.56). Hence, the analysis of the Hall coefficient curve serves as a complementary tool allowing us to determine stepwise for each gate voltage increment the electron and hole densities from the obtained Hall coefficients, as described above (Eq. 4.2). In so doing, also the residual total carrier density at the CNP is determined. Alternatively, in a more elegant approach, the latter can be also derived from fitting the measured Hall coefficient curve by combining the two carrier

6 Only in case of unequal electron and hole mobilities, the Hall coefficient depends on their ratio (Eq. B.21) leading to an asymmetric curve close to the CNP, as we will discuss in the next section. Typically, the band structure and the scattering mechanisms are electron-hole symmetric in graphene. Then, the Hall coefficient for weak magnetic fields is only a function of n_e and n_h.

type Drude model and the channel approximation (Eq. 2.30)[7]. As already shown in Section 2.3.3, Equation 4.1 can then be rewritten as follows

$$R_H(\mu^2 B^2 \ll 1) = -\frac{n_g}{e\, n_t^2} = -\frac{n_g}{e\,(n_{t,0}^2 + n_g^2)} = -\frac{\beta\,(V_g - V_{CNP})}{e\,(n_{t,0}^2 + \beta^2\,(V_g - V_{CNP})^2)}\,. \quad (4.3)$$

This equation involves a minimal[8] set of parameters in order to model the Hall coefficient curve: the residual carrier density at the CNP $n_{t,0}$, the gate voltage corresponding to the CNP V_{CNP}, and the gate factor β. In our practical procedure, we obtain the gate factor already from fitting the single carrier model to the high gate voltage regime, as explained above and demonstrated in Figure 4.1c. We then choose the gate voltage corresponding to the resistivity maximum, $V_g(\rho_{max})$, as the starting value for V_{CNP}. However, we still consider V_{CNP} as a free fitting parameter, because its value might slightly differ in case of the Hall coefficient curve, and already this slight difference may lead to a worse quality fit (see also Section 4.2). Finally, the most important free parameter is the residual total density at the CNP $n_{t,0}$.

Using the fitting routine described here, we obtain good results modeling the experimentally obtained Hall coefficient curves, as demonstrated in panels b and d of Figure 4.1. We emphasize again that the two carrier type Drude model assumes the system to be spatially homogeneous and that the residual density at the CNP $n_{t,0}$ is an ad-hoc parameter introduced by the channel approximation. Hence, the quality of the fit is better when thermally excited carriers prevail over the disorder induced electron hole puddles, so that we are dealing with a quasi-homogeneous system at the CNP. This condition is realized in high quality, hBN-encapsulated graphene devices at sufficiently high temperature, such as for instance room temperature. Bare graphene devices on SiO_2, on the contrary, are still dominated by disorder and the fit quality is worse, as shown in Figure 4.1b. Nevertheless, the fit still allows us to obtain a good estimate of $n_{t,0}$. Refinements of the two carrier type Drude model and an estimation of the error bars are discussed in the following sections. We will demonstrate that applying the simplest form of the two carrier type Drude and the channel approximation in order to describe the measured Hall coefficient curves

7 See also Section 2.2.4 and Figure 2.5f, where we discuss and demonstrate the applicability of the channel approximation with the ad-hoc parameter $n_{t,0}$, for instance by comparing it to the modified DOS model.

8 In the literature, an additional pre-factor $\alpha = r_H\,G$ is sometimes added to Eq. 4.3 [114, 142, 198]. Its use as free fitting parameter is then justified since the Hall scattering factor r_H and the geometrical factor G may differ from unity and are typically unknown. As we already argued in Section 2.3.1, the geometrical factor for our devices with $L/W \geq 3$ is close to one [150]. Moreover, we can also assume the Hall scattering factor to be approximately one [142]. We will further show in Sections 4.2 and 4.3 that the additional pre-factor α does neither improve the fit quality nor influence substantially the other fitting parameters, in particular $n_{t,0}$. Typically, the fitting algorithm returns a value close to one.

leads to reasonable results within these error bars. Therefore, the approach based on Equation 4.3 constitutes our main tool to quickly and reliably analyze our data. It has been also used to collect the device statistics discussed in the next subsection.

Finally, as discussed in detail in Section 2.2.4, the residual total carrier density at the CNP can be modeled as follows

$$n_{t,0}(T) = n_{t,0}(T = 0) + \frac{\pi k_B^2}{3\hbar^2 v_F^2} T^2 \,. \tag{4.4}$$

Assuming that disorder is absent, such that $n_{t,0}(T \to 0) = 0$, and that the Fermi velocity takes on the conventional value $v_F = 10^6\,\mathrm{m/s}$, we can calculate the thermal limit of the Hall coefficient curve at room temperature (green line in Figure 4.1b, see also Ref. 142). Surprisingly, the experimentally obtained curve for hBN-encapsulated graphene reveals substantially larger Hall coefficient values. It suggests that the assumption of a trivial linear band structure of graphene is incorrect. This constitutes one of the main findings of this work and will be elaborated in Section 4.3. Finally, Equation 4.3 also implies a relation between the residual total density at the CNP and the maximum absolute value of the Hall coefficient. When $|n_g| = n_{t,0}$, the maximum absolute Hall coefficient is given by

$$|R_{H,max}| = \frac{1}{2\,e\,n_{t,0}} \,. \tag{4.5}$$

Hence, optimizing the device quality, that means reducing the disorder related part of $n_{t,0}$, will result in much larger maximum Hall coefficients.

4.1.3. Device statistics at room temperature: maximum Hall coefficient and residual density

In this work, we have examined a large number of devices that were fabricated following different routines. These processing steps were discussed thoroughly in the previous chapter. In Section 3.2.2 we already presented the statistics of our devices in terms of their room temperature transport properties in the absence of a magnetic field (see Figure 3.8). Here, we provide complementary statistics focusing on the (weak field) magneto-transport properties at room temperature. As aforementioned and demonstrated in Figure 4.1b, the Hall coefficient curves for all device types are identical at high gate voltages and pass exactly through zero at the CNP. However, they differ substantially in the crossover region, especially in terms of the maximum Hall coefficient that can be reached. Since we are interested in applying graphene for Hall sensing, the maximum sensitivity constitutes a key figure of merit. Hence,

we plotted the measured room temperature values for each device in Figure 4.2 on the y-axis, distinguishing different fabrication routines. The error bars account for the maximum value (on the hole side) and the absolute value of the minimum (on the electron side). In the ideal case of electron-hole symmetry, those values should be identical. Moreover, the Hall coefficient curves allow us to make a good estimate of the residual carrier density at the CNP using the above discussed simplest form of the two carrier type Drude model and the channel approximation. We applied this approach to all studied devices and took the resulting $n_{t,0}$ values at room temperature as the abscissae (filled symbols in Figure 4.2). In addition, we also derived the $n_{t,0}$ values from the corresponding conductivity versus density diagrams at zero magnetic field using the semi-manual, graphical analysis which we explained in detail in Section 3.2.2 and Figure 3.8a. The data is replotted as open symbols employing those $n_{t,0}$ values as abscissae (still using the directly measured maximum Hall coefficient values as ordinates). Typically, the graphical determination of the residual density values at the CNP constitutes a rougher estimate than the fitting of the Hall coefficient curve, since also the mobility might change with density (while the method assumes the mobility to be density-independent). Despite this issue, all data points representing different device types, obtained either from the Hall coefficient curve fitting (filled symbols) or from the graphical analysis (open symbols), clearly follow the relation[9] between the maximum Hall coefficient and the residual density at the CNP that we expect from our simple model (Eq. 4.5, black line in the figure). Our results are further supported by other data obtained on bare graphene on SiO_2 devices (see Ref. 114, not included in our figure). In this work, we prove the validity of the correlation between the maximum Hall coefficient and the residual carrier density over a much wider range of $n_{t,0}$ values and for different device types.

In agreement with Figure 3.8, Figure 4.2 further reveals unambiguously the beneficial effect of our optimized device fabrication routine that we developed in this work (see previous chapter). Uncovered graphene devices on SiO_2 show the worst properties, i.e. the smallest values of the maximum Hall coefficient, due to the high level of extrinsic disorder. The typical values are below 2000 V/AT at room temperature, in agreement with other reports [114, 142, 348]. The encapsulation of graphene by hBN then enhances the performance. However, only the high temperature annealing step at 500 °C and the AFM ironing step, together, unleash the optimum device properties[10]. On average, we obtain maximum Hall coefficient values on the order

9 We emphasize that, in case when the residual density values are derived graphically from the double-logarithmic diagram, the relation is even confirmed by independently determined quantities.

10 We emphasize that the difference in performance can only be revealed by electrical (magneto-) transport measurements. Pre-characterization of hBN-encapsulated devices by means of Raman

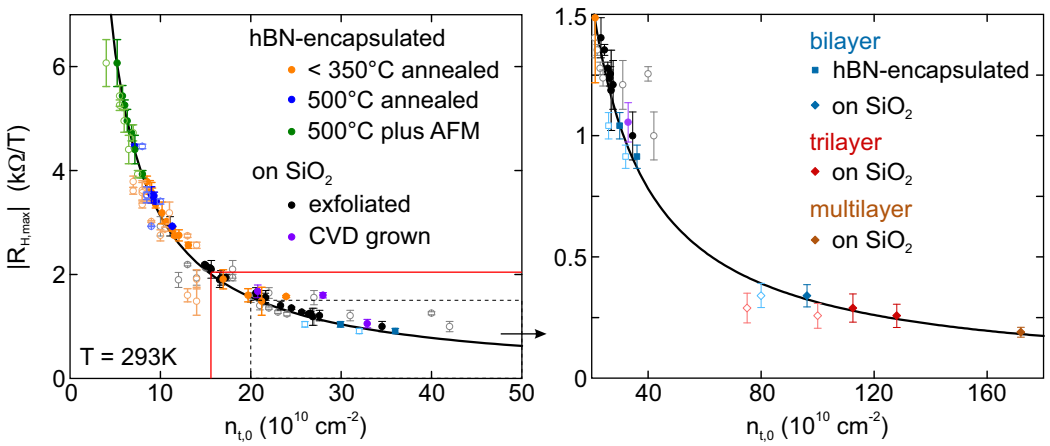

Figure 4.2. Device statistics of the measured maximum Hall coefficient at room temperature and the determined residual carrier density at the CNP. Each data point corresponds to an individual sample. We distinguish different fabrication routines by the respective color of the symbol. The error bar of the maximum Hall coefficient reflects the maximum and the absolute value of the minimum Hall coefficient on the hole and electron side, respectively. The residual carrier density is either determined by fitting the simplest form of the two carrier type Drude model within the channel approximation to the measured Hall coefficient versus gate voltage traces (filled symbols), or by the double-logarithmic method (open symbols) which only requires a measurement of the resistivity. The black line is calculated from $R_{H,max}^{-1} = 2\,e\,n_{t,0}$. The red lines mark the expected thermal limits when $v_F = 10^6$ m/s. The right panel shows a zoom-in (and extension) of the area defined by dashed lines in the left panel. It also contains data on bilayer and few-layer devices. The maximum Hall coefficient at room temperature quickly drops with increasing layer number.

of 5000 V/AT. Our very best device even reaches 6000 V/AT. Similar values were also reported in Ref. 349 studying, however, only two hBN-encapsulated devices (5700 V/AT and 4100 V/AT). Our fabrication routine allows obtaining values of about 5000 V/AT reliably, which we demonstrate for several devices that also exhibit a smaller variation in performance. Furthermore, we emphasize that, for these devices of highest quality, the CNP coincides almost perfectly with zero gate voltage (that means $V_{CNP} \approx 0$). Hence, the maximum Hall coefficient values can be reached by applying a voltage of only few volts (depending on the dielectric thickness even below one volt). This is of great importance when considering the use of graphene based Hall sensors in battery powered devices.

spectroscopy does not indicate any significant difference when comparing different annealing temperatures as well as before and after applying the AFM ironing, as shown in Figure 3.6 for different annealing temperatures. The Raman technique is just not sensitive enough to the small variation of an already quite low disorder level which, however, is reflected substantially in the electrical measurements (see also our work in Ref. 183).

Referring to Section 4.1.2, we point out that thermally activated carriers should limit the maximum Hall coefficient and the $n_{t,0}$ value at room temperature to the area enclosed by the red lines in Figure 4.2, when we assume a linear band structure and $v_F = 10^6\,m/s$. Our data clearly suggests that the violation of this limit is systematic and observed for almost all hBN-encapsulated devices. We attempt to address this issue in Section 4.3. Even though our approach, based on the simplest form of the two carrier type Drude model, does not reproduce the measured electron-hole asymmetries in case of some devices, the resulting error in $n_{t,0}$ can be considered as small and does not alter the conclusion made here (see also Section 4.2). Moreover, the much larger than expected Hall coefficients were observed directly in the experiments and are independent of the model.

Finally, we also included data obtained on bilayer and multilayer devices in Figure 4.2. Due to their parabolic band structure, the number of thermally activated carriers is much larger (see Section 4.4 for details). Therefore, the room temperature performance of the best, hBN-encapsulated bilayer devices does not even exceed the worst monolayer devices on SiO_2. This justifies why we focus on monolayer devices in this work.

4.1.4. Side note on the determination of the mobility

Before we continue the discussion about magneto-transport in the remaining sections of this chapter, we give an example of the different ways, that are utilized in the literature, to extract the mobility and discuss when the mobility term makes sense at all. Generally, in order to determine the charge carrier mobility from the measured longitudinal resistivity (conductivity) in the absence of a magnetic field, a precise knowledge of the carrier density, or the total carrier density in case of the two carrier type regime, is needed (see also Section 3.2.2). As shown in the previous subsections, the measurement of the Hall coefficient as a function of applied gate voltage now enables us to determine independently the gate factor β as well as the residual total carrier density at the CNP $n_{t,0}$ (when using the two carrier type Drude model within the channel approximation)[11]. These two parameters can then be used to calculate the gate controlled average density, $n_g(V_g) = \beta(V_g - V_{CNP})$, and the total carrier density, $n_t(V_g)^2 = n_g(V_g)^2 + n_{t,0}^2$ as function of applied gate voltage.

11 In doing so, the residual density is obtained from a second, independent measurement, in contrast to the approach involving a semi-manual graphical analysis of the double-logarithmic diagram of the conductivity (at zero magnetic field) versus average density, as shown in panel a of Figure 3.8a.

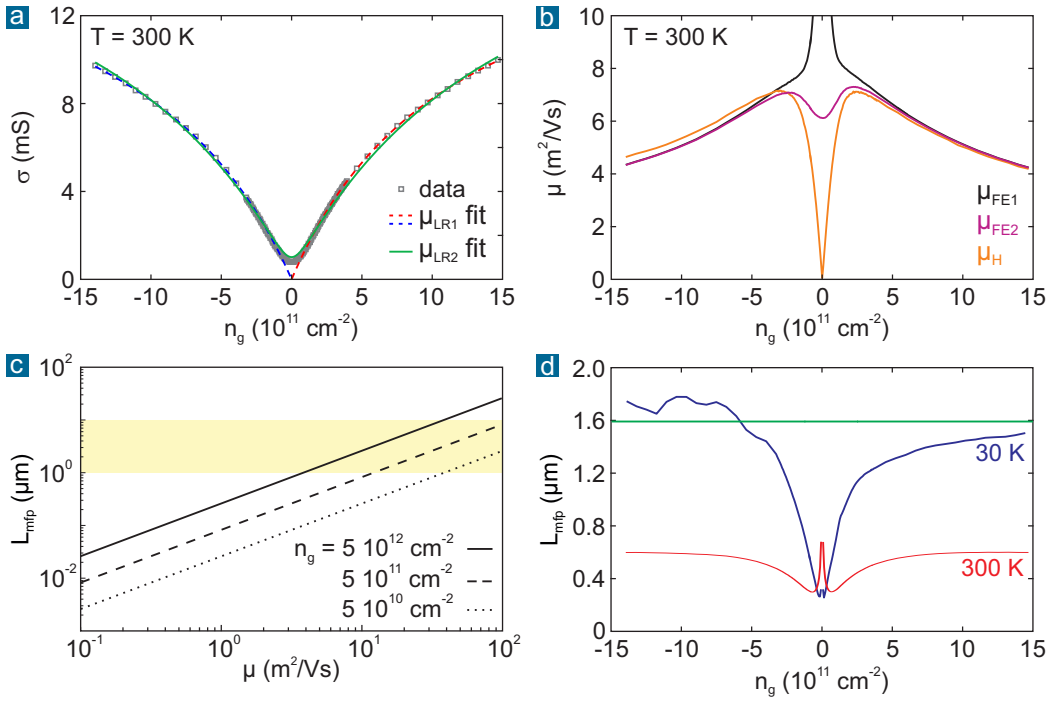

Figure 4.3. Different ways to define the mobility as well as the mean free path that we obtained from the measured longitudinal conductivity (at zero magnetic field). We discuss here the same hBN-encapsulated device as in Figures 2.6c and 3.8a. The exact gate factor and the residual total carrier density at the CNP have been obtained independently by fitting the corresponding Hall coefficient curves (not shown) using the two carrier type Drude model within channel approximation (see main text). a: Fits of the measured conductivity versus the gate controlled average density considering long range type and short range type disorder in the single (Eq. 2.51) and two carrier type picture (Eq. 4.6). b: Field effect mobility in the single (Eq. 2.38) and two carrier type picture (Eq. 2.58), as well as Hall mobility (Eq. 2.42) combining the measurements of the longitudinal conductivity and the Hall coefficient (see main text). c: Theoretical relation between mean free path and mobility for three different densities according to the Boltzmann transport theory for the single carrier type regime (Eq. 2.49). The typical distance between two longitudinal voltage probes of our devices is marked by the yellow area. d: Mean free path obtained from the measured conductivity assuming the single carrier type model (Eq. 2.48). The green line corresponds to the "ballistic" limit (as explained in the main text).

We consider here the same hBN-encapsulated device for which we already presented conductivity versus density diagrams at zero magnetic field in Figures 2.6c and 3.8a. Figure 4.3a shows the measured conductivity at room temperature as a function of n_g, however, using the independently determined gate factor β in order to precisely convert the gate voltage to density. The excellent fit of the two carrier type

Drude model to the corresponding Hall coefficient curve is not shown here explicitly (for an example of another device see Figure 4.1b). From the latter we find that $n_{t,0}(300\,K) = 8.0 \cdot 10^{10}\,cm^{-2}$. Connecting to Section 2.3, different mobility values, following the most common definitions in the literature, can then be extracted directly from the measured quantities (the conductivity and the Hall coefficient). They are plotted in Figure 4.3b as a function of the gate controlled average density n_g:

- Assuming that only a single carrier type is present, the field effect mobility $\mu_{FE,1} = \sigma/e\,|n_g|$ (Eq. 2.38) and the Hall mobility $\mu_H = |R_H|\,\sigma$ (Eq. 2.42) can be defined. At high density, e.g. $n_g = 10^{12}\,cm^{-2}$, both definitions give us the same mobility value on the order of $50,000\,cm^2/Vs$. However, as already discussed in Section 2.3.1, the mobility values start to differ when $k_BT \ll E_F$ no longer holds, that means near the CNP. μ_H drops to zero since R_H crosses through zero, while the conductivity remains finite. On the contrary, $\mu_{FE,1}$ diverges because $n_g \to 0$ in the single carrier type picture.

- Taking into account the coexistence of electrons and holes at the CNP, the total carrier density replaces n_g in the formula of the Drude conductivity. Using the channel approximation and the independently determined residual total carrier density at the CNP (from the Hall coefficient curve fit), we can define a two carrier type field effect mobility $\mu_{FE,2}$ according to Equation 2.58 in Section 2.3.3. As shown in Figure 4.3b, the mobility then remains finite at the CNP (which we referred to as μ_{CNP} in Eq. 3.4 in Section 3.2.2)[12].

Quite often, also another mobility definition is utilized in the literature. As described in detail in Section 2.3.1, it is based on the Boltzmann theory considering both long and short range type disorder in graphene:

- Equation 2.51 describes the density dependence of the conductivity in the single carrier regime using two parameters: The density-independent mobility $\mu_{LR,1}$ related to the long range disorder contribution (such as charged impurities) and the density-independent resistivity $\rho_{SR,1}$ related to the short range disorder contribution (such as acoustic phonons at higher temperatures). The blue (red) dashed line in panel a of Figure 4.3 corresponds to the fit of the hole (electron) side in the high density, single carrier regime. Averaging the electron and hole side values, we obtain $\mu_{LR,1} = 92,600\,cm^2/Vs$ and $\rho_{SR,1} = 54.7\,\Omega$.

12 The dip in the two carrier type mobility at the CNP might be either related to the simplicity of our model, using the channel approximation, or it could be also explained by enhanced optical phonon scattering (see also Section 2.3.2 and Figure 2.9a).

- Again, we can account for the coexistence of electrons and hole at the CNP by incorporating the channel approximation in Equation 2.51. It then reads

$$\frac{1}{\sigma} = \frac{1}{e\,\mu_{\mathrm{LR},2}\sqrt{n_{t,0}^2 + n_g^2}} + \rho_{\mathrm{SR},2}\,. \tag{4.6}$$

  Using $n_{t,0}$ obtained from the fitting of the independently measured Hall coefficient curve, the green line in Figure 4.3a represents the best description of the measured conductivity. The resulting fit parameters are $\mu_{\mathrm{LR},2} = 82,800\,\mathrm{cm}^2/\mathrm{Vs}$ and $\rho_{\mathrm{SR},2} = 47.4\,\Omega$, that means roughly ten percent smaller than when considering the single carrier model.

We emphasize that the obtained long range disorder mobility values are notably larger than the field effect or Hall mobility values we have derived previously. This has to be taken into account when comparing reported mobility values in the literature. As already concluded in Section 3.2.2, we use two reference values for the mobility in this work (that means as a figure of merit for the device quality): The single carrier type field effect mobility at $n_g = 10^{12}\,\mathrm{cm}^{-2}$ (which is equivalent to the Hall or two carrier type field effect mobility at this high density) and the mobility at the CNP (which corresponds to the two carrier type field effect mobility for $n_g = 0$). In both cases, the gate factor and the residual carrier density at the CNP, that are necessary to calculate the mobility values, are obtained independently from the corresponding Hall coefficient curves.

Since hBN-encapsulated graphene devices exhibit very high mobility values even at room temperature (and no matter which definition is eventually used) but are typically limited to few micrometer in size, the system can easily enter the ballistic transport regime. If the mean free path L_{mfp} exceeds the distance between the longitudinal voltage probes, the diffusive picture and hence all previous mobility definitions no longer hold and the conductivity is ultimately limited by boundary scattering. This would result in a square root dependence of the conductivity on the density, as already discussed in Section 2.4 (Eq. 2.71) and illustrated in Figure 2.6c and 3.8a. In Section 2.3.1 we derived the relation between the mean free path and the mobility (or measured conductivity) in the single carrier regime (see Eq. 2.48 and 2.49). Using these equations, we have calculated theoretically the mean free path L_{mfp} as a function of mobility for three different densities, as shown in Figure 4.3c. Ideally, the mobility of graphene at high density is limited only by acoustic phonon scattering (at elevated temperatures, such as room temperature). At room temperature, it is expected to be about $100,000\,\mathrm{cm}^2/\mathrm{Vs}$ ($10\,\mathrm{m}^2/\mathrm{Vs}$). Then, the distance between the voltage probes measuring the longitudinal conductivity should be at least a few micrometer.

The above discussed device exhibits the shortest distance between the voltage probes compared to all in more detail studied devices in this work (about two micrometer). Since the sample is also not perfectly clean, and thus not limited by acoustic phonon scattering alone, the mobility values at room temperature are slightly lower. Accordingly, the mean free path, which we derived directly from the measured conductivity using the single carrier type model (Eq. 2.48), remains below one micrometer (see Figure 4.3d). However, when the temperature is lowered, phonon scattering is suppressed and the conductivity at higher densities approaches the ballistic limit (as already shown in Figure 2.6c and 3.8a for the same device). This is also demonstrated in Figure 4.3d, when we use the measured conductivity at low temperature to calculate the mean free path. At higher densities it approaches the green line which we obtained by calculating the ballistic conductivity for the device's channel width (Eq. 2.71) and converting it to a "false mean free path" according to the semiclassical model (Eq. 2.48). This "false mean free path" is density-independent and determined solely by the width of the channel. Near the CNP, however, the extracted mean free path values are clearly smaller which demonstrates that transport is diffusive at lower densities. Hence, the extraction of mobility values at higher densities requires particular carefulness and a check for plausibility.

The conclusion of this side note is that, for high quality hBN-encapsulated graphene devices, the extracted mobility values not only depend on the utilized definition but also are reasonable, without any restriction to the density range, only at higher temperatures (e.g. room temperature). Hence, the previously presented device statistics of the mobility values at room temperature is accurate (see Figure 3.8 in Section 3.2.2). In the following sections which involve also low temperature measurements, we will not extract mobility values but report directly measurable transport quantities, such as the Hall coefficient or the longitudinal resistivity (or conductivity) at zero magnetic field. We will also focus on the low density regime near the CNP, where the system can be considered as diffusive at all temperatures. Possible phenomena beyond the diffusive picture, occurring at higher densities and low temperature, are discussed only briefly in Section 4.6.

4.2. Electron hole asymmetries

The in Section 4.1.2 suggested simplest form of the two carrier type Drude model, based on a minimal set of fitting parameters, assumes electron hole symmetry. However, this is valid only in the ideal case. Several Hall coefficient curves which we

obtained on real devices exhibit more or less distinct asymmetries between the electron and hole side. In particular, we find those asymmetries to be more evident for devices of lower quality, e.g. on SiO_2, over the whole temperature range. However, they also occur for high quality, hBN-encapsulated devices mostly at lower temperatures. The observed asymmetries do not appear to follow a common trend but are found to be sample-specific. The asymmetry especially manifests in a difference between the absolute values of the maximum and minimum Hall coefficient (see also the error bars in Figure 4.2). Moreover, after converting the Hall coefficients into the density n_{RH} using the previously described single carrier model, these values do not always follow a carrier type independent linear trend in the high density regime. Instead, for some samples, it seems that the gate factor is slightly different on the electron and hole side, respectively. In other words, the $n_{RH}(V_g)$ curve appears to possess a kink at $V_g = V_{CNP}$, as shown in Figure 4.5f.

These experimental findings raise two important questions: how can we cope with this electron hole asymmetry and how imprecise are the results for the main fitting parameters when we employ the simplest, symmetric form of the two carrier type Drude model? In order to address in particular the second question we can compare the fit parameters obtained when applying the simplest form of the two carrier type Drude model to the same Hall coefficient curve in four different ways, e.g. by considering different gate voltage ranges or performing a symmetrization of the original data before the fitting[13]. This is illustrated exemplarily in Figure 4.4 for a Hall coefficient curve obtained on a hBN-encapsulated device at room temperature exhibiting a small degree of asymmetry:

- Case 1 (panel a): standard procedure as introduced in Section 4.1.2 using the complete gate voltage range and the original (asymmetric) data
- Case 2 (panel b): like case 1 but we symmetrized[14] the original experimental data before running the fit procedure (green squares in panel b)

13 However, we emphasize that, in all cases, we still employ the simplest form of the two carrier type Drude model in combination with the channel approximation, and the gate factor β is determined separately from the $n_{RH}(V_g)$ dependence (see Eq. 4.3 and Section 4.1.2).

14 The symmetrization of the data is performed as follows. We estimate V_{CNP} from the maximum resistivity and use this gate voltage value as the starting point. Then, the original Hall coefficient data is interpolated in order to obtain equal gate voltage steps on both sides with respect to this central point. For each increment around V_{CNP}, the absolute Hall coefficient values on the electron and hole side are added and then divided by two. Finally, this new, averaged Hall coefficient value is assigned to the effective gate voltage step on the hole side (and accordingly, its negative value on the electron side). Panel b of Figure 4.4 shows the symmetrized Hall coefficient curve (green squares).

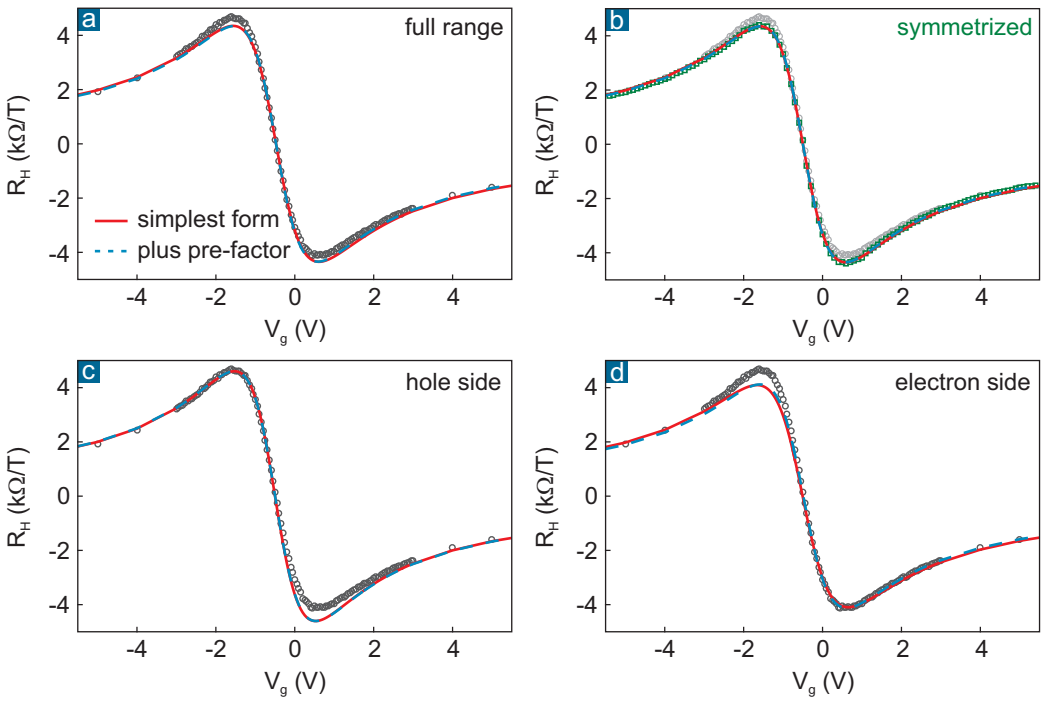

Figure 4.4. Asymmetries between the electron and hole side of the measured Hall coefficient curve. All panels show the same data (gray circles), obtained on a hBN-encapsulated device at room temperature. Here, we focus on a more narrow gate voltage range around the CNP revealing the asymmetries. In order to obtain an estimate of the error bars of the main fitting parameters, e.g. the residual density at the CNP $n_{t,0}$, we apply the simplest form of the two carrier type Drude model given by Eq. 4.3 (red solid line) to the whole gate voltage range (panel a), to the symmetrized data (b), as well as to the hole or electron side only (c,d). See main text and Table 4.1 for details. Moreover, we also demonstrate that taking into account an additional pre-factor α in Eq. 4.3 results in almost identical results (blue dashed line).

- Case 3 and 4 (panels c and d, respectively): we can also model the measured Hall coefficient curve considering only the hole side ($V_g < V_{CNP}$) or the electron side ($V_g > V_{CNP}$) for the fitting algorithm

For all cases, the resulting fit parameter values are listed in Table 4.1. In particular, we obtain four different values of the residual total density at the CNP, $n_{t,0}$. This allows us to calculate an average value and its standard deviation in order to estimate the accuracy. This will be of great importance in the next section when we discuss the temperature dependence.

case	simplest form (Eq. 4.3)	added pre-factor α
case 1: full range and original, asymmetric data	$n_{t,0} = 7.18$ $V_{CNP} = -0.48$	$n_{t,0} = 7.02$ $V_{CNP} = -0.48$ $\alpha = 0.98$
case 2: full range and symmetrized data	$n_{t,0} = 7.17$ $V_{CNP} = -0.50$	$n_{t,0} = 7.07$ $V_{CNP} = -0.50$ $\alpha = 0.99$
case 3: hole side and original, asymmetric data	$n_{t,0} = 6.78$ $V_{CNP} = -0.50$	$n_{t,0} = 6.81$ $V_{CNP} = -0.50$ $\alpha = 1.01$
case 4: electron side and original, asymmetric data	$n_{t,0} = 7.62$ $V_{CNP} = -0.51$	$n_{t,0} = 7.25$ $V_{CNP} = -0.49$ $\alpha = 0.96$

Table 4.1. Fit parameters corresponding to Figure 4.4. $n_{t,0}$ is given in units of $10^{10}\,cm^{-2}$ and V_{CNP} in V. In all cases, the same gate factor $\beta = 6.67 \cdot 10^{10}\,cm^{-2}/V$ was used. We determined it separately by fitting the $n_{RH}(V_g)$ dependence in the single carrier regime, as already described in Section 4.1.2.

Furthermore, we also evaluated the impact of introducing an additional pre-factor α in the model (Eq. 4.3), since this has been proposed and used by other groups (see Refs. 114, 142, 198 and Section 4.1.2). Again, we determine the gate factor β separately from the $n_{RH}(V_g)$ dependence. Therefore, with the addition of the pre-factor the model offers three free parameters (α, V_{CNP}, and $n_{t,0}$) to describe the Hall coefficient curve as opposed to our standard model with only two free parameters. Figure 4.4 clearly shows that the quality of the fit is not improved and both approaches lead to almost identical results (red solid line and blue dashed line). For the example covered here, the values of α are very close to one, and the $n_{t,0}$ values vary only slightly. As a matter of fact, they lie within the error bar extracted from the four different ways of fitting using the simplest form without pre-factor (as described previously). For some of our devices, the value of α may deviate somewhat more from unity (typically between 0.9 and 1.1). Hence, we will routinely monitor the impact of the additional pre-factor on the resulting $n_{t,0}$ values for each device and temperature step studied in the next section. If $n_{t,0}$ changes notably, this variation will be taken into account when stating the error bar of $n_{t,0}$. As already mentioned in Section 4.1.2, when measuring devices with geometrical factors close to one, the introduction of an additional pre-factor is unnecessary.

Two carrier type Drude model with entire parameter set

Instead of restricting us to the simplified two carrier type Drude model, which assumes electron hole symmetry, we can also use the full list of possible parameters for the fitting. This encompasses the parameter $b = \mu_e/\mu_h$, describing the ratio between electron and hole mobility, as well as the Hall scattering factors $r_{e,h}$ on the electron and hole side. In Section 2.3.3 we already derived the complete formula for the Hall coefficient (Eq. 2.63). In combination with the channel approximation (Eq. 2.31), we obtained Equation 2.64 which is repeated here for convenience:

$$R_H = \frac{2}{e} \frac{(r_h - r_e\, b^2)\sqrt{n_{t,0}^2 + \beta^2 (V_g - V_{CNP})^2} - (r_h + r_e\, b^2)\, \beta\, (V_g - V_{CNP})}{\left[(b+1)\sqrt{n_{t,0}^2 + \beta^2 (V_g - V_{CNP})^2} + (b-1)\, \beta\, (V_g - V_{CNP})\right]^2}\,. \quad (4.7)$$

The equation provides six fit parameters: β, V_{CNP}, $n_{t,0}$, b, r_e, and r_h. Panel a of Figure 4.5 illustrates how the new parameters (b, r_e, and r_h) influence the shape of the Hall coefficient curve. In our calculation we used typical values for the remaining parameters, i.e. $V_{CNP} = 0\,V$, $\beta = 7.2 \cdot 10^{10}\,cm^{-2}/V$, and $n_{t,0} = 1.6 \cdot 10^{11} cm^{-2}$ (which corresponds to the thermally excited total density at the CNP at 300 K assuming $v_F = 10^6\,m/s$). The red line corresponds to the symmetric case we considered so far, that means $b = r_e = r_h = 1$. When we now choose $b \neq 1$ but still set $r_e = r_h = 1$, the Hall coefficient curve becomes asymmetric around the CNP but coincides with the symmetric case in the high gate voltage, single carrier regime (green line). On the contrary, when $r_e \neq r_h \neq 1$ but $b = 1$, the symmetry is broken over the whole gate voltage range (orange line). This circumstance becomes especially obvious when the Hall coefficient values are converted into density values using the single carrier model. This results in $n_{RH} = n_g/r_{e,h} = \beta\, V_g/r_{e,h}$. The inset of panel a distinguishes the case when $r_e = r_h = 1$ (magenta line) and the case when $r_e = 0.8$ and $r_e = 1.2$. The latter leads to a kink in the $n_{RH}(V_g)$ dependence (blue and red line representing the hole and electron side, respectively).

It should also be emphasized that the zero crossing of the Hall coefficient curve no longer occurs exactly when $V_g = V_{CNP}$, if one of the parameters b, r_e or r_h is not equal to one (see panel a). As mentioned in Section 2.3.3, R_H becomes zero when the enumerator in Equation 4.7 is zero. Obviously, this constitutes a more complex condition than just $V_g = V_{CNP}$, which also depends on the asymmetry parameters b, r_e and r_h.

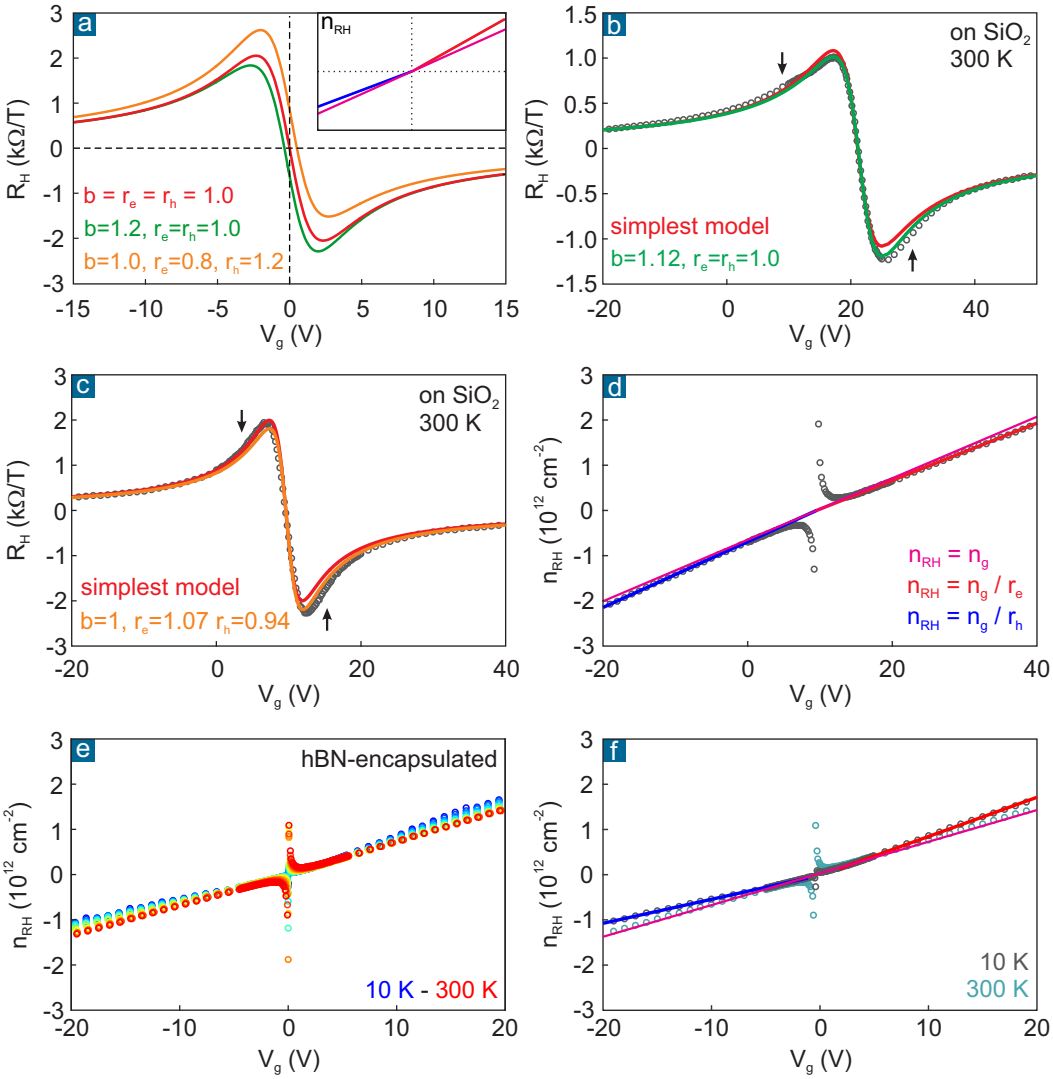

Figure 4.5. Modeling of the observed electron hole asymmetry by means of the two carrier type Drude model and the channel approximation using the complete set of fit parameters. We take into account different electron and hole mobilities ($b = \mu_e/\mu_h$) as well as different Hall scattering factors for the electron and hole side (r_e, r_h). Data obtained on three different devices are shown: Two graphene devices on SiO_2 (shown panel b and panels c,d) and one hBN-encapsulated device (panels e,f). See main text for the discussion.

Finally, it can be shown that not all six parameters in Equation 4.7 can be simultaneously considered as free to fit the experimental Hall coefficient curves[15].

15 As already discussed in Section 2.3.3, changing for instance β by an arbitrary factor m leads to the exact same solution when $n_{t,0}$ and r_e, r_h are scaled by the same factor m. As a consequence, the optimization algorithm will not be able to find a singular minimum.

Therefore, in analogy to our previous procedure using the simplest form of the two carrier type Drude model, we first analyze the extracted density values n_{RH} as a function of gate voltage in the high density, single carrier regime. However, now we determine individual gate factors on the hole side, β_h, and on the electron side, β_e, by selecting the respective data points (to allow for the asymmetric behavior). We then calculate the global gate factor β as their average[16] value, and accordingly, the Hall scattering factors through $r_{e,h} = \beta/\beta_{e,h}$. Panels d and f of Figure 4.5 illustrate this procedure for two different devices: graphene on SiO_2 at room temperature (panel d) and hBN-encapsulated graphene at low temperature (panel f). For the hBN-encapsulated device, panel e demonstrates clearly that the kink in the $n_{RH}(V_g)$ curve develops as the temperature decreases. The averaged gate factor at low temperature, represented by the magenta line, agrees well with the room temperature data plotted additionally in panel f (turquois circles), which indicates only a weak temperature dependence. As a matter of fact, different Hall scattering factors for the electron and hole side are necessary in order to describe the kink in the low temperature data (gray circles and blue and red lines).

Following the described procedure, we obtain β (averaged) and r_e, r_h for each device (and temperature step in the next section). These are then kept fixed in Equation 4.7 when fitting the Hall coefficient curve over the whole gate voltage range, while V_{CNP}, $n_{t,0}$, and b, constitute free parameters. This approach improves the fit quality notably in comparison to the simplest form of the two carrier type Drude model. Panels b and c demonstrate this exemplarily for two different graphene devices on SiO_2 at room temperature. For the first sample (panel b) we find the $n_{RH}(V_g)$ dependence to be electron hole symmetric and hence set $r_e = r_h = 1$. The asymmetry around the CNP is then nicely captured by $b = 1.12$. This value also conforms to good approximation to the ratio of the electron and hole mobility values we can estimate from the longitudinal resistivity (at zero magnetic field). However, the impact on the value of the residual density $n_{t,0}$ is marginal. We obtain $n_{t,0} = 28.37 \cdot 10^{10}\,cm^{-2}$ using the full parameter set, and $n_{t,0} = 28.86 \cdot 10^{10}\,cm^{-2}$ using the simplest form of the two carrier type Drude model. The reason for this is that introducing the parameter b appears to have the same effect as symmetrizing the raw data. For the second device (panel c), the converted density values do not follow perfectly a common linear trend for both carrier types in the single carrier regime (panel d). In order to account for this, we introduce $r_e = 1.07$ and $r_h = 0.94$. Then, the fit to the experimental Hall coefficient curve returns $b \approx 1$ and $n_{t,0} = 15.69 \cdot 10^{10}\,cm^{-2}$ (orange line). Although the quality of the fit is improved, the residual density is again almost identical to the value we obtain using the simplest model, $n_{t,0} = 15.63 \cdot 10^{10}\,cm^{-2}$. Obviously,

16 Typically, this averaged gate factor still has a value comparable to the one we would obtain when considering data points on both the electron and hole side together (magenta lines).

the introduction of r_e and r_h and slight changes of the gate factor β do not influence significantly the result for the residual density value[17]. These additional parameters are less relevant in the low density regime near the CNP, while for $n_{t,0}$ this regime is.

All in all, we find that the results for the residual density in all our different approaches are robust. The simplest form of the two carrier type Drude model, applied to either both or only one side of the two carrier regimes, as well as the two carrier type Drude model using the entire set of parameters yield nearly identical results. In the next section, we will apply all these approaches routinely to each measured Hall coefficient in order to get a reliable estimate on the accuracy of the residual density value. This will allow us to make a firm statement about thermally excited carriers at the CNP.

We would like to remind that the two carrier type Drude model assumes a spatially homogeneous system. However, electron and hole puddles form at lower densities at low temperature. For samples with a higher level of disorder these puddles even appear at room temperature. This may explain the minor but yet discernible deviations between the two carrier type Drude model and the experimental Hall coefficient curves. These deviations arise in particular in the crossover region between the CNP and the single carrier regime at high density. The arrows in panels b and c of Figure 4.5 point to the deviations in case of the two graphene devices on SiO_2. As discussed in Section 2.3.3, it has been shown, for devices with high disorder level or at low temperature, that the effective medium theory can describe better the density dependence of the conductivity as well as the magneto-resistance curves in the crossover region than the two carrier type Drude model. Unfortunately, an explicit calculation of the Hall coefficient curve based on the EMT model has to our best knowledge not yet been published[18].

The physical origin of the observed asymmetry remains partially unclear. The introduction of the parameter b, the ratio between the electron and hole mobility, is

17 This is also supported by the fact that the obtained $n_{t,0}$ values do not change notably when we modify our procedure for obtaining the gate factor β. In our standard approach, the latter is extracted from the $n_{RH}(V_g)$ curve restricting us to the single carrier regime. If we take all data points of the n_{RH} curve into account, including the CNP region, the derived gate factor varies and the fit quality becomes worse, but the resulting value for $n_{t,0}$ remains quite stable.

18 The calculations conducted in Ref. 110 may serve as a starting point providing explicit expressions for σ_{xx}^{EMT} and σ_{xy}^{EMT}. Moreover, the EMT theory has so far been applied only to symmetrized data, e.g. magneto-resistance curves [110]. Hence, it is not capable to account for electron hole asymmetries. Finally, a theoretical concept describing the crossover from the disorder dominated regime to the spatially homogeneous regime with a majority of thermally excited carriers is lacking.

reasonable since it conforms with longitudinal transport data in the absence of a magnetic field. In contrast, we cannot explain the notable difference between the electron and hole side at high density in the single carrier regime that we observe on some devices. Even though this asymmetry can be described properly using carrier type specific Hall scattering factors ($r_e \neq r_h \neq 1$) in the two carrier type Drude model, it is unlikely that these parameters have a proper physical meaning (but just constitute additional degrees of freedom for the fit). According to its definition (Eq. 2.41), the Hall scattering factor depends on the dominant scattering mechanism in the material [150, 152]. This would imply that the dominant scattering mechanisms on the electron and hole side are different. Moreover, neither do all devices show this effect nor do those who show it reveal a common trend. Often, the effect is found to be enhanced as the temperature is lowered. One may argue that the field effect gating depends on the polarity or the dielectric material. However, the effect is observed both for the implemented silicon back gate using SiO_2 (and hBN) as dielectric as well for graphitic back gate configurations using only hBN as the dielectric. To conclude, dedicated experiments are needed to achieve a better understanding of the effect's origin. However, as demonstrated above, the impact on the low density regime near the CNP is small and the residual density value is barely affected. Since our work focuses on this region, we refrain from any further investigation of the effect.

4.3. Temperature dependence

For several devices we also studied the temperature dependence of the Hall coefficient curve over a wide temperature range down to pumped liquid helium. Panel a of Figure 4.6 shows this dependence exemplarily for the hBN-encapsulated device for which we already presented the room temperature data in Figure 4.1. In the high density regime, that means when $E_F \gg k_B T$ over the whole temperature range, all Hall coefficient curves virtually coincide. If at all, small deviations can only be discerned when converting the measured Hall coefficients into densities using the single carrier model, as described in Section 4.1.2. Panel b displays the resulting $n_{RH}(V_g)$ traces, revealing a slight deviation from the exact linear relation as the temperature is lowered. As discussed above, the emergence of this effect is sample specific (compare also to the other hBN-encapsulated device displayed in panel e of Figure 4.5). Running our standard analysis procedure of Section 4.1.2, we obtain the gate factor β from fitting the $n_{RH}(V_g)$ curve considering only the high gate voltage data points. In doing so for each temperature value, we notice that the gate factor varies also slightly with temperature, as shown in panel f. We emphasize again that taking

into account a temperature dependence of β improves the fit quality of our model to the experimental Hall coefficient curve but does not alter the results for the residual density values at the CNP in any significant manner[19]. Moreover, the physical origin of this weak temperature dependence remains unclear. In the previous section we argued that especially the kink seen in some $n_{RH}(V_g)$ traces might indicate different Hall scattering factors which themselves can be temperature dependent[20]. In addition, one might also need to consider that, at low temperature and high density, the system enters the ballistic transport regime which can influence the Hall voltage measurement. Finally, the dielectric constant of the gate dielectric material itself, i.e. SiO_2 and hBN, exhibits a temperature dependence. All in all, this may explain the sample-specific and non-monotonic but typically small temperature dependence of the gate factor. In our analysis, uncertainties arising from the latter are well considered by our error estimate.

In the vicinity of the CNP, the Hall coefficient curves change drastically with temperature. The Hall coefficient extrema increase by an order of magnitude as the temperature is lowered, reaching values up to $25,000\,V/AT$ (see also panel c of Figure 4.6). Their distance to the CNP, in terms of effective gate voltage that needs to be applied in order to reach these sensitivity extrema, becomes steadily smaller (panel d). This results in a much steeper slope of the Hall coefficient crossing through zero at the CNP. Moreover, for hBN-encapsulated devices, the gate voltage value corresponding to the CNP, V_{CNP}, is almost a constant of temperature. The V_{CNP} values derived independently from the maximum resistivity at zero magnetic field as well as from fitting the simplest two carrier type Drude model to the measured Hall coefficient curves, using V_{CNP} as adjustable parameter, are plotted in panel e.

The data displayed in Figure 4.6 is representative for all hBN-encapsulated devices for which we measured the whole temperature range. However, in case of uncovered graphene devices on SiO_2, the temperature dependence of the Hall coefficient curves is typically subject to extrinsic influences which impede a clear analysis[21].

19 For instance, this was also tested by assuming a temperature independent, fixed value of the gate factor. The resulting $n_{t,0}$ values remain within the stated error bars.

20 A similar explanation based on temperature dependent Hall scattering factors has also been invoked in Ref. 350 in order to describe the observed change in the gate voltage dependence of the density with temperature for MoS_2 devices.

21 The initial doping level leading to a non-zero V_{CNP} value stems from charge transfer to (polar) adsorbates and other contaminants, and is not inherent to graphene itself (see also Section 3.2.2 and Ref. 240). The adsorption and desorption rates are clearly a function of temperature. Moreover, the possible movement of polar adsorbates is hampered at low temperature. All this can explain the stronger temperature dependence of V_{CNP} and other odd effects in the field effect gating for non-encapsulated devices.

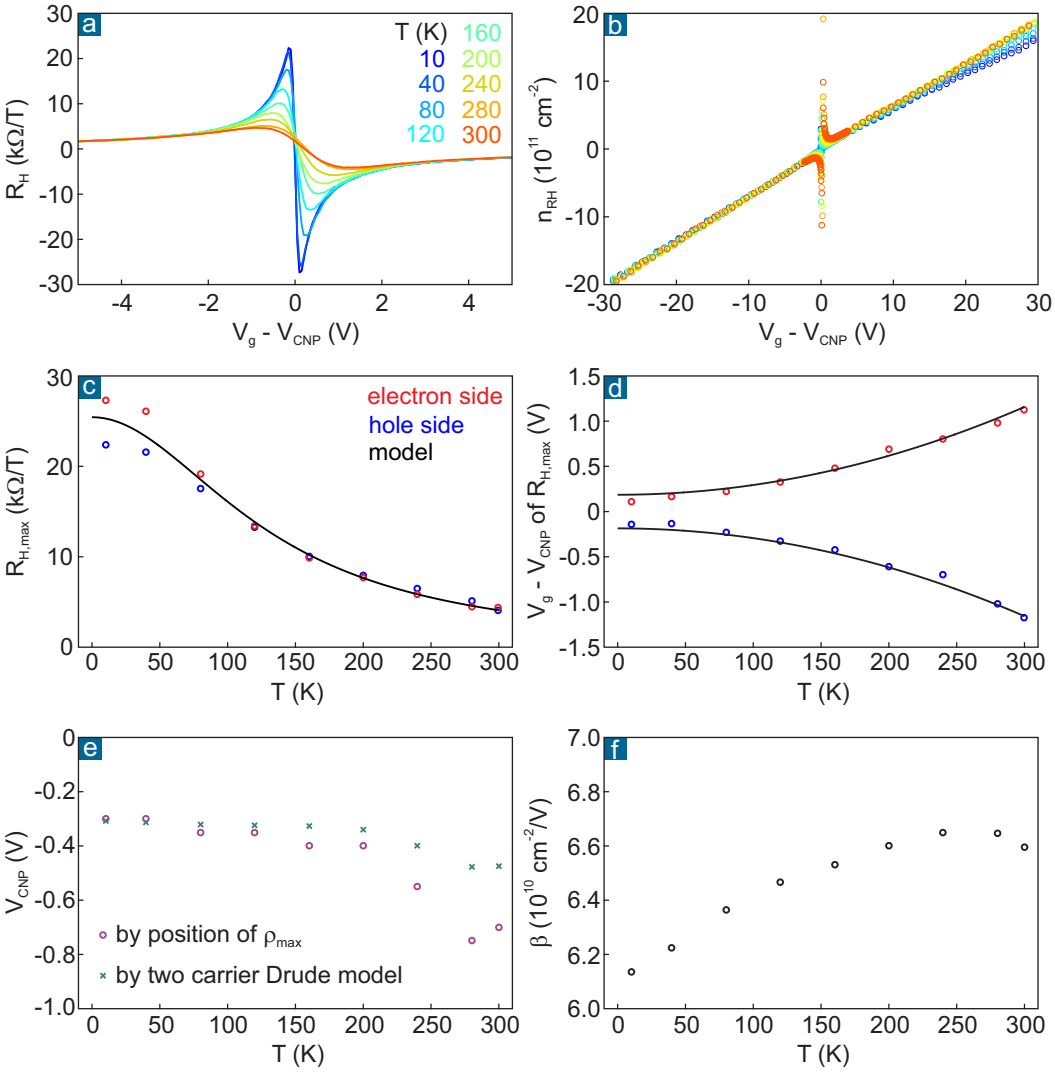

Figure 4.6. Exemplary temperature dependence of the Hall coefficient curve as a function of gate voltage for hBN-encapsulated graphene (panel a). b: Data from panel a converted into density values using the single carrier model. c: Experimentally obtained Hall coefficient extrema on the electron and hole side as a function of temperature. The model is discussed in Section 4.3.1. d: Temperature dependent position of the Hall coefficient extrema with respect to the CNP. e: Shift of V_{CNP}, defined as the position of ρ_{max} or obtained as fitting parameter describing the $R_H(V_g)$ dependence (gate voltage steps are 0.1 V in the experiment). f: Gate factor β as a function of temperature, obtained from panel b.

For instance, we often observe a much stronger and even non-monotonic shift of V_{CNP}. Moreover, the Hall coefficient curves can become more asymmetric and even

deformed in an odd way that cannot be accounted for properly by our model. We further experienced problems in terms of increasing contact resistance and noise at lower temperatures. Finally, the disorder level and so the amplitude of the electron hole puddles is considerably larger in these non-encapsulated devices than in their hBN-encapsulated counterparts (see also Sections 2.2.3 and 3.2). Therefore, the thermal excitation of carriers at the CNP will not yet prevail over the disorder induced density fluctuations at room temperature. The temperature dependence of the Hall coefficient curves is notably weaker. As a consequence, only two of our devices on SiO_2, the best ones in terms of initial doping level and density inhomogeneity, provided reasonable, unambiguous information on the temperature dependence.

Finally, we proceed by performing our analysis routine, which we discussed in detail in Sections 4.1.2 and 4.2, on the Hall coefficient curves of all studied devices that show a clear temperature dependence. We both employ the simplest form of the two carrier type Drude model (with three parameters) as well as the Drude model with entire parameter set in order to obtain an estimate of the accuracy of the extracted residual density values (see Figure 4.7 and Table 4.2). When using Equation 4.4 to model the measured temperature dependence of the residual carrier density at the CNP, $n_{t,0}(T)$, a clear deviation between our experimental data and the expected thermal excitation of carriers, governed by a linear energy dispersion with the conventional Fermi velocity value of $v_F = 10^6\,m/s$, is seen. This agrees with our earlier finding that, for many devices, the maximum Hall coefficients at room temperature are larger than what we anticipated (see Section 4.1.3). This issue will be addressed thoroughly in the following.

4.3.1. Thermal excitation of charge carriers at the CNP

How can we explain this large discrepancy between the maximum Hall coefficient in experiment and the theoretically predicted value? The discrepancy becomes particularly apparent when we compare the experimentally determined residual density value at room temperature, $\approx 7 \cdot 10^{10}\,cm^{-2}$, and the expected value for disorder-free graphene with $v_F = 10^6\,m/s$, $\approx 16 \cdot 10^{10}\,cm^{-2}$ (see Section 2.2.2). As discussed in Section 2.1.7, there is strong evidence from other types of experiments that many body interactions become important at low density in high quality devices with a very low level of disorder. It has been asserted that the band structure and the Fermi velocity are renormalized at low energies. A theoretical description of the thermal excitation of charge carriers considering such renormalization effects have to the best of our knowledge not been developed, i.e. carrier-carrier interactions have been

neglected so far. None of the current theoretical models have simultaneously addressed many body interactions as well as the effect of disorder, causing the system to break up into electron hole puddles, in a common framework.

In the absence of such an advanced model, we can approximate the complex physics in our devices by means of an average, energy independent Fermi velocity that is larger than the conventional value of the tight binding model (which is based on a single particle picture)[22]. In the following we use this promising and conclusive approach[23] to analyze the temperature dependence of the residual density that we obtained from the measured Hall coefficient curves. The two parameters in Equation 4.4, $n_{t,0}(T = 0)$ and v_F, are adjusted in order to properly describe the data points. The quality of the fit is excellent, as demonstrated in Figure 4.7 for the exemplary hBN-encapsulated device. We then can use these fit parameters and Equation 4.4 as input for the simplest two carrier type Drude model in order to calculate the Hall coefficient extrema (via Equation 4.5) and their positions with respect to the CNP (given by the condition $n_g = n_{t,0}$) as a function of temperature. Again, panels c and d of Figure 4.6 show that the agreement between this approach and the data points, obtained directly from the measured curves without any further data processing, is convincing.

In order to estimate the accuracy of our results, we apply Equation 4.4 to all $n_{t,0}(T)$ data sets we obtained by using the different fitting procedures based on the two carrier type Drude model, as described above. For the exemplary hBN-encapsulated device, Table 4.2 lists the $n_{t,0}(T = 0)$ and v_F values for the different data sets. We then calculate the mean value and standard deviation of $n_{t,0}(T = 0)$ and v_F for each studied device. One could now pose the question whether the effect of renormalization, i.e. change in v_F, is larger when the strength of disorder is smaller.

22 Recall that the approximation of the complex physics by means of a larger but still density independent, or energy independent, Fermi velocity value is also motivated by ARPES measurements and tunneling spectroscopy experiments analyzing the Landau levels (see Refs. 55, 63 and Figure 2.4 in Section 2.1.7). Both experiments allow the investigation of the spectral function for fixed density over a wide energy range (while in transport measurements, the energy is fixed by the density). The spectral function is found to have a steeper slope but to remain linear for any low density value that was investigated in the experiments.

23 We can exclude that density dependent Hall scattering factors are the reason for the larger Hall coefficients. This would imply that we underestimate the residual density. Similar density values are however also obtained from the conductivity versus density diagrams in the absence of a magnetic field (where Hall scattering factors do not play a role). For instance, the data in the literature, which we discuss below, is typically derived in that way. We also considered the formation of bound excitons (electron hole pairs) as a possible cause for the smaller free carrier density values we see in our experiments and when analyzing the data in the literature. However, all other experiments in the literature, using different measurement techniques, point to a renormalization of the Fermi velocity as the most likely cause.

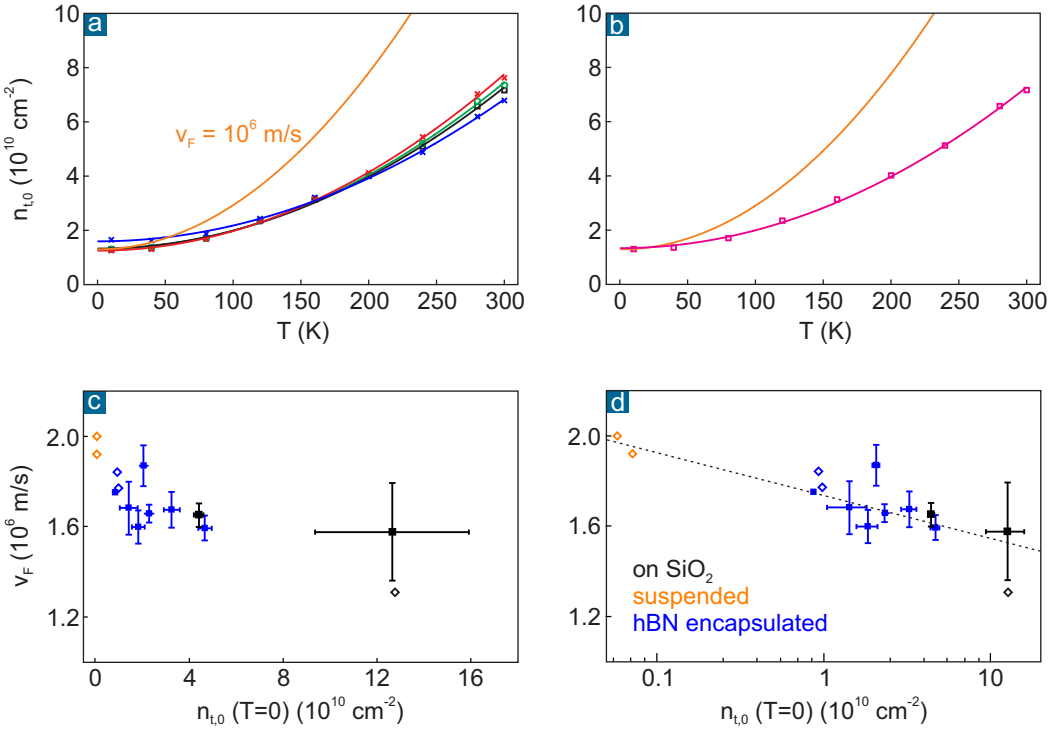

Figure 4.7. Temperature dependence of the residual total carrier density at the CNP $n_{t,0}$. The residual density values are obtained from fitting the two carrier type Drude model to the measured Hall coefficient curves. a: The simplest form of the model is used to describe the curves in Figure 4.6a. We compare the results for $n_{t,0}$ considering the entire gate voltage (density) range (black symbols) and only the electron or hole side (red or blue symbols) for the fitting. Moreover, the results of the fit to the symmetrized data are plotted (green symbols). The solid lines correspond to Eq. 4.4 describing the temperature dependence of the residual density at the CNP. The experimental data clearly deviates from the expectation based on the conventional value of the Fermi velocity (10^6 m/s, orange line). Using v_F and $n_{t,0}(T = 0)$ as free parameters, we can nicely reproduce the $n_{t,0}(T)$ dependence (see Table 4.2). b: Same procedure as in panel a but using the two carrier type Drude model with the entire set of parameters (magenta symbols). c: The diagram summarizes for each studied device the v_F and $n_{t,0}(T = 0)$ values that describe best the obtained temperature dependence of $n_{t,0}$. Filled symbols correspond to our data based on the analysis of the Hall coefficient curves by means of the two carrier type Drude model (which allows us to provide error bars). The open symbols represent elsewhere directly reported v_F and $n_{t,0}(T = 0)$ values [82] or were determined by us applying Eq. 4.4 to the $n_{t,0}(T)$ data points provided in other studies [45, 79, 126, 195]. See main text for details. d: Same data as in panel c but x-axis is plotted on a logarithmic scale. Dashed line serves as a guide to the eye.

two carrier type Drude model used for fitting $R_H(V_g)$ in order to obtain the $n_{t,0}(T)$ dependence		fit to the obtained $n_{t,0}(T)$ dependence	
form	density range	$n_{t,0}(T=0)$	v_F
simplest	entire	1.33	1.65
simplest	electron side	1.25	1.58
simplest	hole side	1.56	1.76
simplest	symmetrized data	1.31	1.62
simplest plus α	symmetrized data	0.98	1.58
entire set incl. b, r_e, r_h	entire	1.34	1.65

Table 4.2. Resulting values for v_F (in 10^6 m/s) and $n_{t,0}(T=0)$ (in $10^{10}\,cm^{-2}$) describing best the $n_{t,0}(T)$ dependence for the exemplary device presented in Figure 4.7.

Since $n_{t,0}(T=0)$ constitutes a measure of the disorder strength, we have plotted the extracted Fermi velocity values as a function of $n_{t,0}(T=0)$ in panels c and d of Figure 4.7. Each data point corresponds to one of the studied devices (filled squares, the error bars represent the standard deviations). In total, we consider the results for seven hBN-encapsulated devices and two graphene on SiO_2 devices, which showed a clear temperature dependence of the Hall coefficient curves[24].

Our results clearly demonstrate that the thermal activation of charge carriers at the CNP is governed by a substantially larger Fermi velocity value than the tight binding model predicts. Given the markedness of this effect seen on all device types, it raises the question whether other groups have reported similar observations. Surprisingly, not many studies have provided data on the temperature dependence of (magneto-) transport, at zero or classically low magnetic fields, over a wide temperature range. Mostly, these studies focused on other effects and commented, if at all, only briefly on the peculiar temperature dependence. In some works, larger Fermi velocity values were discussed but the dependence on the disorder level was never investigated. In the following, we list the main findings of these other studies:

- Zhu et al. (Ref. 82) reported on the temperature dependence of Hall coefficient curves obtained on one SiO_2 supported graphene sample between liquid helium and room temperature. Using an approximation of the two carrier type Drude transport theory, they modeled the steep (linear) slope around the CNP and

24 As aforementioned, the temperature dependence of literally dirtier graphene on SiO_2 devices with higher disorder level, e.g. $n_{t,0}(T=0) \approx 5 - 30 \cdot 10^{10}\,cm^{-2}$, is weak and affected by many other effects, and thus difficult to analyze.

obtained the $n_{t,0}(T)$ dependence, in our notation, which they then fitted using Equation 2.35. They found $v_F = 1.3 \cdot 10^6\,m/s$ and the disorder parameter $s = 54\,meV$. This corresponds approximately to $n_{t,0}(T = 0) \approx 13 \cdot 10^{10}\,cm^{-2}$ when using Equation 4.4. Note that in Equation 2.35 v_F and s are not entirely independent parameters, in contrast to $n_{t,0}(T = 0)$ and v_F in Equation 4.4. This may lead to slightly different results for v_F. The result of Zhu and co-workers is plotted in panels c and d of Figure 4.7 (black diamond), along with our data points.

- Farmer et al. (Ref. 351) studied the temperature dependence of the Hall mobility in epitaxial graphene on SiC from liquid helium to room temperature. Using the two carrier type Drude model, they analyzed the temperature dependence of the density values corresponding to the maximum Hall mobility, $n(\mu_{H,max})$. Their data suggest $v_F \approx 1.5 \cdot 10^6\,m/s$. Unfortunately, we cannot convert their data to comply with our approach and a value equivalent to $n_{t,0}(T = 0)$ is missing. Still, their finding also supports the picture of a larger Fermi velocity value.

- Mayorov and co-workers investigated the temperature dependence of the width and the amplitude of the minimum conductivity at the CNP for suspended, current-annealed graphene devices with extremely low disorder (Refs. 126, 195). Based on the two carrier type Drude model and assuming a density independent mobility, they demonstrated that the number of carriers in graphene, $N = \sigma(n_g)/e\mu$, close to the CNP can be approximated by a quadratic function of the thermally excited carrier density. Analyzing their data from helium temperature up to 150 K, the authors provide a $n_{t,0}(T)$ dependence (in our notation). Using the same model as we do (Eq. 4.4), they estimated $v_F \approx 2 \cdot 10^6\,m/s$ and $n_{t,0}(T = 0) \approx 5 \cdot 10^8\,cm^{-2}$. The authors then argued that the cause of the larger Fermi velocity is the renormalization at low energies. However, they did not proceed with a detailed analysis. In Figure 4.7 we included the results we obtained from fitting their $n_{t,0}(T)$ data points that we extracted from their publications (orange diamonds).

- In the supplementary information of Ref. 45, Ponomarenko and co-authors examined the thermal broadening of the resistivity peak at zero magnetic field for a hBN-encapsulated graphene device that exhibits a Moiré superlattice (see also Section 4.4.2). Applying the same data analysis as Mayorov et al. in Refs. 126, 195, they obtained the $n_{t,0}(T)$ dependence between 70 K and 200 K at the h-SNP as well as at the main CNP. We extracted their data points for the main CNP and ran our fitting routine based on Eq. 4.4. Again, we added the result to our figure (one of the blue diamonds).

- Finally, Crossno et al. measured the temperature dependence of longitudinal transport in the absence of a magnetic field for hBN-encapsulated devices. In their work (Ref. 79), which mainly discusses the observation of the breakdown of the Wiedemann-Franz law, they provide $n_{t,0}(T)$ data points. These were obtained from a graphical analysis of the conductivity as a function of density in a double-logarithmic diagram, as we described already in Section 3.2.2. Note that this approach is also based on the two carrier type Drude model and the channel approximation, and assumes a density-independent mobility. We extracted their data points between 4 K and 200 K and employed Eq. 4.4 for fitting (second blue diamond data point in our figure).

All in all, the data of other groups is consistent with our findings. Unambiguously, the thermal excitation of carriers at low densities cannot be described by a single-particle, non-interacting theory, such as the tight binding approximation with $v_F \approx 1 \cdot 10^6$ m/s. Instead, a renormalized energy spectrum has to be taken into account. In particular, we emphasize that these renormalization effects play a crucial role understanding the room temperature properties of high quality graphene devices. Hence, they have also significant implications on application relevant specifications, e.g. the maximum Hall coefficient.

Finally, we wish to address briefly two other issues. First, our approach and all other approaches used in the literature to estimate the residual density at the CNP from (magneto-) transport curves involve the two carrier type Drude model. This model, however, assumes a homogeneous system. Will we obtain similar results, i.e. a larger Fermi velocity, using other models, such as the effective medium theory? Second, should we expect a correlation between $n_{t,0}(T = 0)$ and v_F, what our data in combination with the data in the literature allegedly suggest?

Analysis of the data by means of the effective medium theory

In order to answer the first question, we remind of Section 2.3.3 where we introduced the current state of the effective medium theory in the literature. The model accounts for the spatial inhomogeneity near the CNP which occurs in all devices at low temperature and can still prevail at room temperature for devices with a higher level of disorder. As discussed in Section 2.2.3 and Table 2.1, the disorder temperature scale is about 100 K for $n_{t,0}(T = 0) \approx 10^{10}\,\mathrm{cm}^{-2}$. Below that temperature, it has been shown in Ref. 80 that the temperature dependence of the conductivity versus density curves can be described by the EMT model considering only charged impurity disorder. However, Hall coefficient curves have not yet been described by

means of the EMT theory. While the manuscript of this thesis is being written, we are in contact with the theory group headed by Shaffique Adam (author of Refs. 7, 80, 109, 110). We provided them with one data set comprising the temperature dependence of the conductivity at zero magnetic field as well as the corresponding Hall coefficient curve. In a first approach, Dr. Indra Yudhistira of this group adapted the EMT model and fitted our magneto-transport data (connecting to their work in Ref. 110). First, the conductivity as a function of gate controlled average density was modeled at low temperature from which the root mean square value of the density fluctuations n_{rms} as well as the density-independent mobility μ were obtained (see Section 2.3.3 and Refs. 7, 109). He then assumed this mobility value to be temperature independent and used it as a fixed parameter while fitting the Hall coefficient curves at all temperatures collectively. The n_{rms} value was also considered as a constant of temperature, however, it was used again as adjustable parameter (since its result varies slightly from the $\sigma(n, T \to 0)$ fit)[25]. Under these assumptions, a temperature dependence only arises from the thermal excitation of charge carriers. The basic concept of the EMT theory is that the inhomogeneous system is modeled by homogeneous, disorder-free patches each with a spatial density distribution described by a Gaussian function [7]. Accordingly, the thermal excitation of charge carriers within these patches simply follows Equation 2.22 in Section 2.2.2. That means, it is governed by the density of states and, hence, by the choice of the band structure parameters. Using the Fermi velocity of the tight binding band structure as second, temperature independent fit parameter, Yudhistira found that $v_F = 1.41 \cdot 10^6\,\mathrm{m/s}$ and $n_{rms} = \sqrt{3} n_{t,0} = 2.66 \cdot 10^{10}\,\mathrm{cm}^{-2}$ describe the temperature dependence of the Hall coefficient curves. In conclusion, the three temperature independent fit parameters μ, n_{rms}, and v_F allow reproducing the measured Hall coefficient curves from liquid helium to room temperature nicely (preliminary results not yet shown), except for asymmetries that are not included in the EMT equations. Improvements of the fit quality could also be achieved by taking into account a temperature (and density) dependent mobility (which would require to incorporate a specific microscopic picture of the disorder/scattering mechanism). All in all, the EMT model leads to very similar results than the two carrier type Drude model. Both approaches indicate a notably enhanced Fermi velocity value when one assumes that the band structure remains linear (which constitutes an approximation of the renormalized spectral function).

25 It can be shown that the $\sigma(n)$ dependence is much more sensitive to the mobility value than to the n_{rms} value, while the opposite is the case for the Hall coefficient curve. As a consequence, it is reasonable to use n_{rms} again as free parameter since slight changes have notable impact on the fit quality.

Correlation between the renormalized Fermi velocity and the disorder

As discussed in detail in Chapter 2, the intrinsic Dirac point physics manifests stronger for lower disorder, i.e. smaller carrier density inhomogeneities. If this inhomogeneity is larger, renormalization effects, that scale logarithmically with lowering the energy, should become weaker. In their recently published theoretical study on the role of electron-electron interactions in two-dimensional Dirac fermion systems (Ref. 352), the group of Shaffique Adam uses a combination of nonperturbative numerical and analytical techniques in order to provide a conclusive picture of the Fermi velocity renormalization. In the absence of disorder, they show that the renormalization of the Fermi velocity is a function of the short range component of the Coulomb interaction U (Hubbard model on-site component) and its long range component α_0. The authors then identify two extreme cases. One is controlled by the long range interaction and results in an increase of the Fermi velocity, while the other is dominated by short range interactions and leads to a suppression of the Fermi velocity (then, the interactions also gap the system causing a Mott insulator phase transition). Tang et al. state that graphene Dirac fermions lie somewhere in between these two regimes.

In the second part of their work (Ref. 352), the authors consider realistic graphene devices that cannot reach the assumed condition of exact half filling. In other words, in these devices it is not possible for the Fermi energy to reach the Dirac point globally (at every point of the sample), due to the presence of disorder causing spatial density inhomogeneities. This density inhomogeneity can be characterized by $n_{rms} = \sqrt{3} n_{t,0}(T = 0)$. Based on the latter, they then define the experimental probe energy scale $\Lambda_k = \sqrt{\pi n_{rms}}$. Their study shows that their results are valid for any Dirac fermion system for which U, α_0, and Λ_k can be defined. For realistic graphene, the on-site potential is about $U = 3$. Hence, the interaction-caused enhancement of the Fermi velocity is mainly determined by the dielectric environment, since α_0 depends inversely on κ (Eq. 2.10), and by the strength of disorder, characterized by $n_{rms} = \sqrt{3} n_{t,0}(T = 0)$. Tang et al. also demonstrate by means of a renormalization group flow approach that it is possible to extrapolate their numerical findings, calculated for very high disorder on the order of $n_{rms} \approx 10^{14}\,cm^{-2}$, to the experimental window accessible for very high quality samples with $n_{rms} \approx 10^9 - 10^{11}\,cm^{-2}$ (see our data in Figure 4.7). For instance, their calculation predicts a Fermi velocity value of $v_F = 1.34 \cdot 10^6\,m/s$ for $n_{rms} = 10^{10}\,cm^{-2}$ and $\alpha_0 = 1$. Moreover, the value of the Fermi velocity is found to further increase only slowly. The increase is logarithmic as the disorder decreases. In their computation using $\alpha_0 = 1$, the ratio α/α_0 drops from about 0.8 to about 0.5 as n_{rms} decreases from 10^{14} to $10^8\,cm^{-2}$ (i.e. six orders of magnitude).

In panel d of Figure 4.7 we plotted the experimentally obtained Fermi velocity values versus the residual density at the CNP in the low temperature limit, using a logarithmic scale on the x-axis. We emphasize that the disorder in the investigated devices spans only about two and a half orders of magnitude in contrast to the six orders in the theoretical calculations. A vague logarithmic dependence of v_F on $n_{t,0}(T = 0)$ adumbrates only when we include the two data points obtained on suspended graphene devices in the literature (see dashed line as a guide to the eye). Moreover, we also need to take into account that three different device types are plotted together in the diagram. The theory predicts the strongest enhancement of the Fermi velocity for suspended devices with $\alpha_0 = 2.2$, while the effect is weaker for SiO_2 and especially for hBN-encapsulated graphene with $\alpha_0 < 1$. As a consequence, the weak dependence on $n_{t,0}(T = 0)$ might be also misleading or at least both parameters, κ and $n_{t,0}(T = 0)$, contribute. Likely, both parameters also play a role when comparing devices on SiO_2 and hBN-encapsulated devices. The latter have an order of magnitude smaller disorder level but also a smaller α_0 value due to the encapsulation[26]. Therefore, the enhancement of the Fermi velocity due to reduction of disorder is likely weakened. A key experiment to test the theoretical predictions would require that the (statistical) spread of the disorder level is wider for a certain device type. Ideally, one should be able to tune the disorder level in one device under test over a wide range. Unfortunately, this seems to be extremely difficult and unlikely to be achieved in experiments.

For hBN-encapsulated devices, the statistical variation of the disorder level allows to probe at most one order of magnitude (typically, $5 \cdot 10^9 - 5 \cdot 10^{10}\,cm^{-2}$). Literally dirtier devices on SiO_2 are more difficult to measure since the temperature dependence of the resistance peak broadening and the Hall coefficient curves is much weaker even at room temperature, since the disorder induced carriers still prevail. In addition, other effects such as desorption of adsorbates contribute. To conclude, we can state that the theoretical model supports qualitatively our findings. However, a quantitative analysis is difficult due to the discussed reasons. In general, our analysis suggests larger Fermi velocity values, $1.6 - 1.7$, than predicted by the theory, ≈ 1.3. This might motivate modifications to the theoretical framework. Or, it indicates that our analysis procedure based on the two carrier type Drude model as well as the simplified picture of a still perfectly linear band dispersion with larger Fermi velocity value have to be revised in order to achieve a better quantitative agreement. We note that, in a first approach, the EMT model based fitting of the Hall coefficient curves resulted in a Fermi velocity value of ≈ 1.4.

26 In addition, it should also matter that the dominant source of disorder is different for different device types. Charged impurity disorder dominates in uncovered devices. On the contrary, random strain fluctuations become (more) important in encapsulated devices.

4.3.2. Electrical transport at zero magnetic field

In order to complete the picture, we briefly report on the temperature dependence of the electrical transport in the absence of a magnetic field. This has already been studied more intensively in the literature (see Sections 2.3.2 and 2.3.4). According to used two carrier model, which is based on the Drude Boltzmann theory, the temperature dependence of the longitudinal resistivity curves is affected both by the thermal excitation of carriers as well as by the temperature dependence of the mobility i.e. the scattering mechanism. In contrast, the Hall coefficient curves are determined by the carrier densities only. Hence, both measurements can be combined to calculate the mobility, in particular at the CNP, μ_{CNP}, which we already demonstrated in Sections 3.2.2 and 4.1.4.

Figure 4.8 shows the longitudinal resistivity measurements on the same hBN encapsulated device for which we already studied the Hall coefficient curves in Figure 4.6. The measured resistivity data plotted in panel a can be easily converted into the more meaningful double-logarithmic diagram of the conductivity ($\sigma = 1/\rho$) as a function of gate controlled, average density $\mathrm{n_g}$ (panel b). These density values are calculated using the gate factors we determined from the Hall measurements. The shift of $\mathrm{V_{CNP}}$ with temperature is very small for such high quality devices (as shown in Figure 4.6e,f). Referring to Section 3.2.2 and Figure 3.8a, the graphical analysis can be applied in order to estimate the residual carrier density (assuming the mobility to be density independent in the low density regime). The results are similar to those we obtain by fitting the corresponding Hall coefficient curves. All in all, the double-logarithmic diagram demonstrates nicely the broadening of the plateau at the CNP due to the thermal excitation of carriers. Moreover, the minimum conductivity is larger at higher temperatures than at low temperature because of the additional carriers available for electrical transport.

Alternatively, we can plot directly the measured maximum resistivity at the CNP as a function of temperature (panel c of Figure 4.8). Starting from a finite, not so high value (or plateau) at low (liquid helium) temperature, the maximum resistivity value drops following a power-law temperature dependence before it flattens out at higher temperatures. The power-law trend seems to come to a halt or even reverses around room temperature. We emphasize that none of our hBN-encapsulated devices revealed a strong insulating behavior at low temperature, i.e. even at the CNP the resistance remains low. In theory, this insulating behavior was predicted for devices with extremely low disorder, that means if the disorder induced residual density tends to zero [149]. We also never observed an exponential temperature dependence

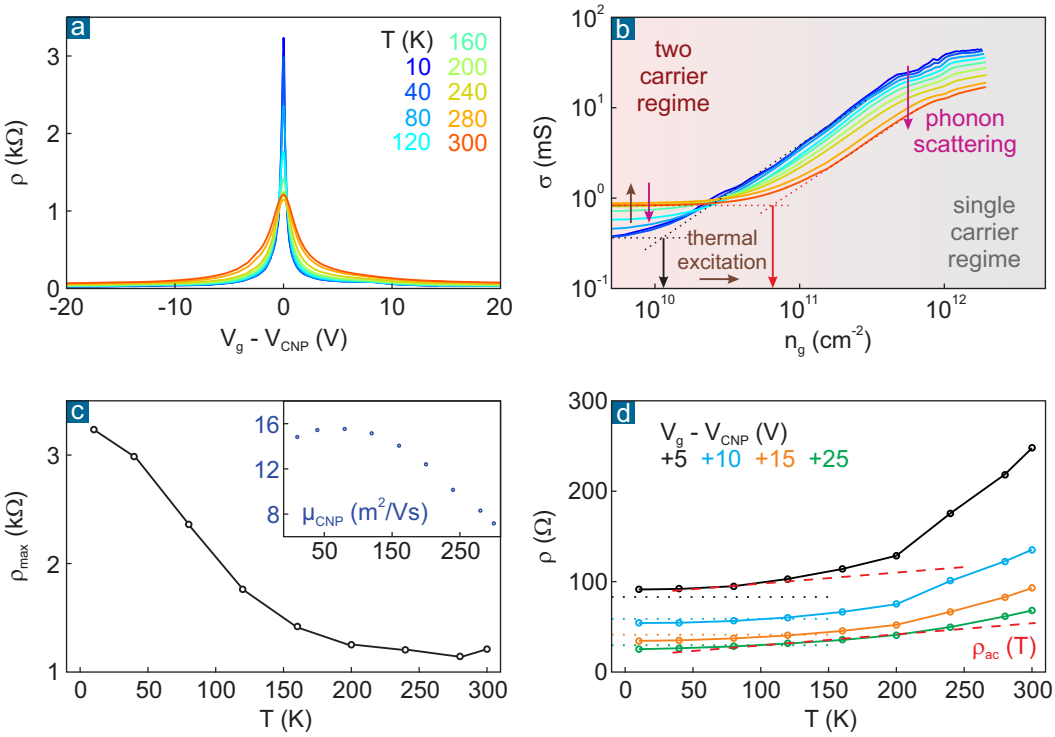

Figure 4.8. Temperature dependence of electrical transport in the absence of magnetic field. a: Longitudinal resistivity as a function of gate voltage for several temperatures. We present exemplary data obtained on the same hBN-encapsulated device for which we showed the Hall coefficient curves in Figure 4.6. b: Double-logarithmic plot of the conductivity over gate controlled, average carrier density (data for the hole side of panel a). See main text for further discussion. c: Maximum resistivity at the CNP as a function of temperature. The inset displays the temperature dependence of the estimated mobility at the CNP. It was calculated by means of Equation 3.4 using the measured ρ_{max} and the $n_{t,0}$ values, obtained from fits to the Hall coefficient curves, as input. d: Temperature dependence of the resistivity for several effective gate voltage values away from the CNP. The dotted lines correspond to the ballistic limit due to boundary scattering which plays a role at low temperature due to the narrow channel width. The expected linear increase of the resistivity due to acoustic phonon scattering, given by Eq. 2.55, is included as dashed red lines (assuming $v_F = 1 \cdot 10^6\,\mathrm{m/s}$ and $D_A = 20\,\mathrm{eV}$). See main text for further discussion.

indicating a band gap formation (as discussed in Section 2.3.4). However, the observed shape of $\rho_{\mathrm{max}}(T)$ nicely conforms to the theoretical picture and to other data measured on suspended devices, as shown in Figure 2.7 in Section 2.3.4 or Ref. 204. The power-law insulating behavior can be explained by the thermal generation of carriers while the saturation and trend reversal around room temperature can be

attributed to phonon scattering. Using the measured maximum resistivity values and the residual carrier density values obtained from the Hall coefficient curves, the mobility at the CNP can be calculated based on the two carrier type Drude model (Eq. 3.4). The results are shown in the inset of panel c and support the above picture. Around 100 K, the mobility starts to decrease steadily. According to the theoretical work of Sohier and coauthors (see Ref. 78 and Section 2.3.2), optical phonon branches are expected to dominate at the CNP. However, we cannot apply their model to our data since renormalization effects would have to be considered in their theoretical framework, requiring calculations beyond the scope of this work.

Finally, the temperature dependence of the resistivity at several other gate voltage values, away from the CNP, are displayed in panel d of Figure 4.8. In the single carrier regime, the thermal generation of carriers is negligible and the change of the resistivity is governed solely by the temperature dependence of the scattering mechanisms, as discussed in Section 2.3.4. Both charged impurity scattering and scattering due to random strain fluctuations exhibit only a weak dependence on temperature, mainly due to the temperature dependent screening function of graphene (see Section 2.3.2 and Ref. 102). Therefore, the temperature dependence will be governed by phonon scattering above a certain threshold temperature. This threshold temperature decreases with increasing density and sample quality. When measuring high quality devices having dimensions of only a few micrometer, we further observe signatures for ballistic transport. Then, boundary scattering determines the measured resistivity value along the channel. For the channel width of our device, we calculated the expected ballistic limits for each studied gate voltage by means of Eq. 2.71 in Section 2.4. The dotted lines in the figure indicate that the resistivity values in the low temperature regime are strongly affected by boundary scattering[27]. The ballistic transport regime then obscures the onset of the linear temperature dependence of the resistivity that is anticipated due to acoustic phonon scattering (dashed red lines, calculated using Eq. 2.55 from Refs. 90, 175). Moreover, when we approach the CNP closer, the measured dependence deviates stronger from the theoretical model predicting a linear dependence and the temperature above which the deviation develops becomes smaller. This agrees with findings reported by other groups that measured supported devices both on SiO_2 and on hBN [10, 39, 90, 161, 188]. Again, the above mentioned theoretical model, that was proposed by Sohier et al. (Ref. 78), provides an explanation. Renormalization effects and scattering with graphene intrinsic optical phonons, that are stronger at lower densities, explain the non-linear temperature dependence in experiment (see Sections 2.3.2 and 2.3.4).

27 The emergence of ballistic transport is further revealed by the wiggles one can clearly notice in the σ(n) curves at low temperature and high density (panel b of Figure 4.8).

4.4. Beyond monolayer graphene

After having gained a proper understanding of the transport in monolayer graphene in the absence and presence of weak magnetic fields, an obvious question is whether we can transfer our models to similar systems such as bilayer or few-layer graphene. We have reported in the sample statistics part of the first section that the maximum Hall coefficients at room temperature are found to decrease substantially with increasing layer thickness. As we will discuss below, the inevitable cause of this behavior is the different band dispersion which leads to a larger number of thermally excited carriers. This highlights the peculiarity of monolayer graphene making it a unique Hall sensor transducer material at room temperature. We also wish to discuss the case when hBN-encapsulated monolayer graphene is well aligned to one of the hBN flakes and a Moiré superlattice emerges (see Section 2.1.6 for the theoretical background). The superlattice leads to cloned Dirac cone features in the transport curves at higher gate voltages which are particularly pronounced at low temperature. We demonstrate that these satellite neutrality points exhibit a different temperature dependence than the original neutrality point (the main CNP). The strong thermal broadening smears out the features at room temperature but still causes asymmetry around the main CNP. Hence, for applications, it is favorable to avoid the alignment between graphene and hexagonal boron nitride.

4.4.1. Hall coefficient with increasing layer thickness

The room temperature Hall coefficient and the resistivity curves (not shown) plotted against the applied gate voltage become flatter and the features broader when the layer number is increased (as shown in Figure 4.9a). Our investigated non-encapsulated devices on SiO_2 also exhibit a notable electron-hole asymmetry in the Hall coefficient curve which leads to a shift of the $R_H = 0$ crossover with respect to the V_{CNP} value which we determine by the resistivity peak. We attribute this observation to the comparatively strong initial (hole) doping we have found in these devices which indicates the presence of extrinsic dopants. However, regardless of the asymmetry which leads to slightly larger Hall coefficients on the electron side, the maximum Hall coefficients at room temperature do not exceed $300\,\Omega/T$. This is by almost an order of magnitude smaller than the routinely measured values for monolayer graphene on SiO_2 devices which are subject to the same device processing steps (compare to Figure 4.1b and Figure 4.2). Motivated by the improvements of monolayer graphene's device quality through hBN-encapsulation, e.g. in terms of

maximum Hall coefficients reaching up to 6000 Ω/T at 300K, we have also examined a hBN-encapsulated bilayer sample (presented in Figure 4.9). Similar to the monolayer case, all main characteristics improve substantially. The initial doping is reduced to almost zero, the mobility increases[28], and the Hall coefficient curve becomes more symmetric with a maximum of about 1000 Ω/T. Although this corresponds to a three-fold improvement, these values are at most comparable to the worst monolayer graphene devices, e.g. in case of CVD grown and wet transferred devices on SiO_2. This clearly does not favor the use of more than a single graphene layer for magnetic field sensing applications when the sensitivity is considered as decisive parameter.

In order to explain these observations, we invoke two possible reasons, interlayer screening effects and the modification of the band structure with increasing layer number. Due to the semimetallic properties of each graphene layer, the field effect gating becomes less efficient due to electrostatic interlayer screening, as reported recently in Ref.354. This leads to a complex charge density distribution in multilayer graphene which is not addressed by our simple two carrier type Drude model which presumes the system to be homogeneous also in the vertical direction. The same argument holds for adsorbates acting as dopants on the sample surface. In sum, we argue that this likely explains the occurrence of electron-hole asymmetry and why our model exhibits a fit of increasingly worse quality when the layer number is increased[29]. As a consequence, the obtained residual densities values, which act as fitting parameter within the channel approximation, constitute only a rough estimate. Our conclusion, that more than one layer graphene is not favorable for magnetic field sensing, is based on the small Hall coefficients measured. It is not altered but supported by the screening problem which is inherent to the system and thus cannot be addressed technically.

Hence, in the following we focus on bilayer[30] graphene which constitutes the next-simplest system for which the electrostatic gating still works sufficiently. Our model (green dotted line) reproduces the measured Hall coefficient curve (dashed blue line)

28 The mobility at the CNP is found to be about $20,000\,cm^2/Vs$ at room temperature (slightly increasing with carrier density). It then rises almost linearly with temperature reaching $40,000\,cm^2/Vs$ at low temperature. In general, the obtained mobility values are smaller in the bilayer than in the monolayer case. Due to the larger residual carrier densities at the CNP, however, the maximum resistivity of the bilayer sample remains comparable to the monolayer.

29 When the layer thickness grows, even the derived densities from the Hall measurements ($n_{RH} = 1/eR_H$) at high gate voltages are no longer properly described by a common gate factor.

30 Note, our argumentation, that the modified band structure results in stronger thermal activation of carriers which limits substantially the maximum Hall coefficients, applies also to thicker layers [82].

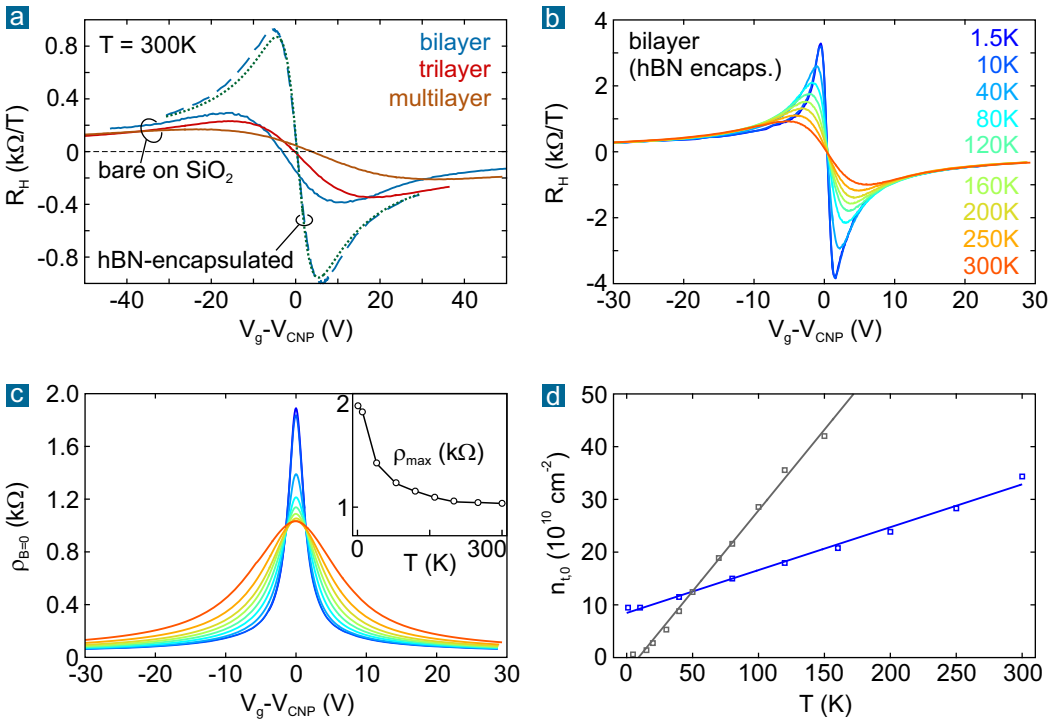

Figure 4.9. Hall coefficient with increasing layer thickness and temperature dependence of bilayer graphene. a: Characteristic Hall coefficient curves versus applied gate voltage for bilayer, trilayer, and multilayer graphene (about 5-6 layers) on SiO_2 (solid lines) and for hBN-encapsulated bilayer graphene (dashed line). The green dotted line represents the fit by our model. Measurements were performed at room temperature and R_H was derived from fitting the Hall resistivity within $\pm 50\,\mathrm{mT}$ magnetic field range. For the hBN-encapsulated bilayer sample, the temperature dependence of the Hall coefficient curves and of the resistivity curves at zero magnetic field is shown in panel b and c, respectively. d: Temperature dependence of the residual total carrier density at the CNP. The blue data points are derived from our hBN-encapsulated bilayer by fitting the Hall coefficient curves. The linear dependence is captured by $m^* = 0.008 m_e$ and $n_{t,0}(0) = 8.5 \cdot 10^{10}\,\mathrm{cm}^{-2}$. For reference, the gray data points are taken from Ref. 353 and are described best by $m^* = 0.031 m_e$.

not as perfect as in the monolayer case but still adequately enough (see Figure 4.9a). For smaller densities, e.g. below $n_g = 10^{12}\,\mathrm{cm}^{-2}$, the tight binding band structure of bilayer graphene can be described using the parabolic approximation which results in a constant density of states, $D_0 = 2m^*/\pi\hbar^2$ [97, 109, 111]. In analogy to Section

2.2.2, the total carrier density at the CNP caused by thermal activation is then given by

$$n_{t,0}(T) = 2\, n_e(\mu = 0, T) = 2\, \ln(2)\, D_0\, k_B\, T \approx 10\, m_{pre}\, T \;\; [\text{in } 10^{10}\, cm^{-2}]\,. \quad (4.8)$$

Here, we have used $m^* = m_{pre} \cdot m_e$ for the effective mass introducing the dimensionless pre-factor m_{pre} which in the following will serve as fitting parameter. In literature, $m_{pre} = 0.033$ is reported as typical value at high densities (e.g. in Refs. 97, 111, 355, 356). Then, the thermally activated, residual total carrier density at the CNP amounts to about $10^{12}\, cm^{-2}$ at room temperature (300K). Using the two carrier type Drude model within the channel approximation, we estimate the maximum Hall coefficient to be slightly above $300\, \Omega/T$ which agrees well with our data obtained on the non-encapsulated bilayer device on SiO_2 and with another report on the same device type [82]. However, in case of our hBN-encapsulated bilayer sample we find the measured Hall coefficients at room temperature to exceed the expected values reaching up to $1000\, \Omega/T$ (in full analogy to the monolayer case).

The temperature dependence of the Hall coefficient and of the resistivity is presented in Figure 4.9b and c, respectively. By fitting our two carrier type Drude model to the measured Hall coefficient curves, we can determine the residual density at the CNP as a function of temperature (blue data points in Figure 4.9d). The expected linear[31] function (Equation 4.8) fits well for the parameters $m_{pre} = 0.008$ and $n_{t,0}(0) = 8.5 \cdot 10^{10}\, cm^{-2}$. The result is astonishing in two ways. First, the determined effective mass is about a factor of four smaller than the expected and commonly reported value[32]. This implies that the thermal activation of carriers is suppressed by the same factor. Again, the analogy to the monolayer samples is apparent. In the latter case, a larger Fermi velocity is necessary to explain the experimentally found smaller pre-factor of the quadratic temperature dependence (note, the pre-factor scales with $1/v_F^2$). Second, the disorder induced residual density at the CNP, which we have defined as the $T \to 0$ limit and used here as second fitting parameter, is quite large and comparable to the typical residual density of

31 An interaction-driven reconstruction of the band structure in bilayer graphene has been reported to occur below 20 K in ultra-clean, suspended devices [195]. The reconstruction leads to a suppression of the DOS which manifests in a much slower decrease of the residual carrier density. Above that critical temperature, however, the complex model quickly becomes indistinguishable from the linear temperature dependence which we take here as a basis.

32 Note, we have employed different variations of the two carrier type Drude model, from the simplest one with a minimum number of fitting parameters to the advanced one involving Hall scattering factors and different electron/hole mobilities, but still obtain similar $n_{t,0}$ values. We also have analyzed the measured resistivity curves and determined $n_{t,0}$ as the crossover point to the minimum conductivity at the CNP (in a double-logarithmic diagram of the conductivity). This analysis then results in $m_{pre} = 0.01$ which still is substantially smaller than the expected value.

hBN-encapsulated monolayer graphene at room temperature (that means including thermal activation). As a consequence, not even the maximum Hall coefficients of bilayer graphene at low temperature exceed the room temperature performance of monolayer graphene.

Two recent studies support our finding that the effective mass of bilayer graphene is not a constant but is suppressed for smaller densities [355, 356]. Both groups investigated bilayer graphene on hBN [356] or even hBN-encapsulated bilayer graphene [355]. By analyzing the temperature dependence of the resistivity curves up to 50 K, Cobaleda and coworkers [355] report a suppression of the effective mass down to $m^* = 0.009\,m_e$ at the CNP. The authors of Ref. 356 examine carefully Shubnikov de Haas oscillations in order to deduce the density dependence of the effective electron and hole masses. The effective masses are found to decrease with smaller densities. They explain their observation not only by the inclusion of electron-electron interaction but also stress the importance of spatial Coulomb potential fluctuations (the electron hole puddle landscape) which is argued to cause a partial cancellation of the effective mass. On the other hand, a well-known work on the thermal broadening of ultra clean, suspended bilayer graphene disagrees with our results [195]. The authors derive the carrier density at the CNP from a fit to the conductivity in a narrow range around the CNP (for temperatures up to 150 K). Their data, multiplied by two to adapt to our $n_{t,0}$ definition, is replotted as gray points in Figure 4.9d. In agreement with the statement in their publication, we find their data to be captured well by $m^* = 0.031\,m_e$ (using Equation 4.8). The difference between their and our results (and the mentioned results of other work on hBN-supported bilayer [355, 356]) is puzzling. In particular, because the same authors also present data on the thermal broadening of suspended monolayer graphene which requires large Fermi velocities of up to $2 \cdot 10^6$ m/s to reproduce the experimental results (both in the same work [195] and in another work [126] using the same analysis procedure). Further experiments may help to understand the cause of the opposing results (for example, whether the hBN support itself or whether the larger disorder induced fluctuations in hBN-supported or encapsulated bilayer than in suspended devices play a role).

4.4.2. Moiré effect on hexagonal boron nitride

When the crystallographic planes of graphene and hexagonal boron nitride are well aligned, a Moiré superlattice emerges (see Section 2.1.6 for the theoretical background). The weak periodic potential induces Dirac point replicas, also called secondary neutrality points (SNP), near the edges of the superlattice Brillouin zone

which can be accessed by conventional electrostatic gating. High magnetic field measurements of such graphene superlattice systems have led to the first experimental observation of Hofstadter's butterfly and fractal quantum hall effects [45, 47, 50, 52]. In the following, we briefly report how the Moiré superlattice influences the weak magnetic field transport properties focusing on the thermal broadening of the hole-side satellite peak and the main CNP, also referred to as original neutrality point (ONP). The temperature dependence of the Hall coefficient and the resistivity curves of our hBN-encapsulated monolayer graphene sample is shown in panels a and b of Figure 4.10. The Hall coefficients at low temperature (10 K) and at room temperature (280 K) are further converted to carrier densities assuming the trivial single carrier formula ($n_{RH} = 1/eR_H$) and plotted in Figure 4.10c. We find the characteristics of the hole-side SNP to be much more distinct than on the electron side which, ultimately, can be traced back to the relative position of the carbon atoms with respect to the boron or nitride atoms causing an asymmetry in the DOS (see Section 2.1.6 for details). Since the features on the electron-side are distinguishable only at low temperature, we omit here the temperature dependence discussion of the electron-side SNP (similar findings regarding the eSNP were also made in Ref. 45).

At low temperature, the maximum resistivity at the hole-side SNP exceeds its counterpart at the ONP[33] but then quickly decreases and the peak width broadens strongly with increasing temperature. On the contrary, the maximum Hall coefficients around the hSNP are substantially smaller than for the ONP and the curves around both SNP are strongly asymmetric. Thus, applying the two carrier type Drude model for the Hall coefficient to fit our data near the hSNP is critical because only few measured Hall coefficients in a narrow gate voltage range around the V_{CNP} equivalent of the hSNP are predominantly determined by the hole-side Dirac point physics. For example, note the deviation of the measured data from the expected n_{RH} values (green line in Figure 4.10c) using the gate factor β which properly describes the n_{RH} values around the ONP. Hence, we extract the residual total carrier density at the hSNP by means of its definition as the crossover point to the conductivity plateau (when plotting the conductivity over density in a double-logarithmic diagram). Figure 4.10d shows the so-determined $n_{t,0}$ values at the hSNP as a function of temperature. The data points clearly follow a quadratic temperature dependence which proves that the band dispersion near the hSNP is

33 We emphasize that we do not observe a strongly insulating behavior at the ONP in case of our sample, in agreement with Refs. 45, 50. Such behavior, however, was sometimes reported for the same type of graphene-hBN superlattice systems and attributed to a band gap formation at the main CNP [47, 52]. The band gaps are then explained by an interplay between structural and electronic properties (as discussed in Ref. 51).

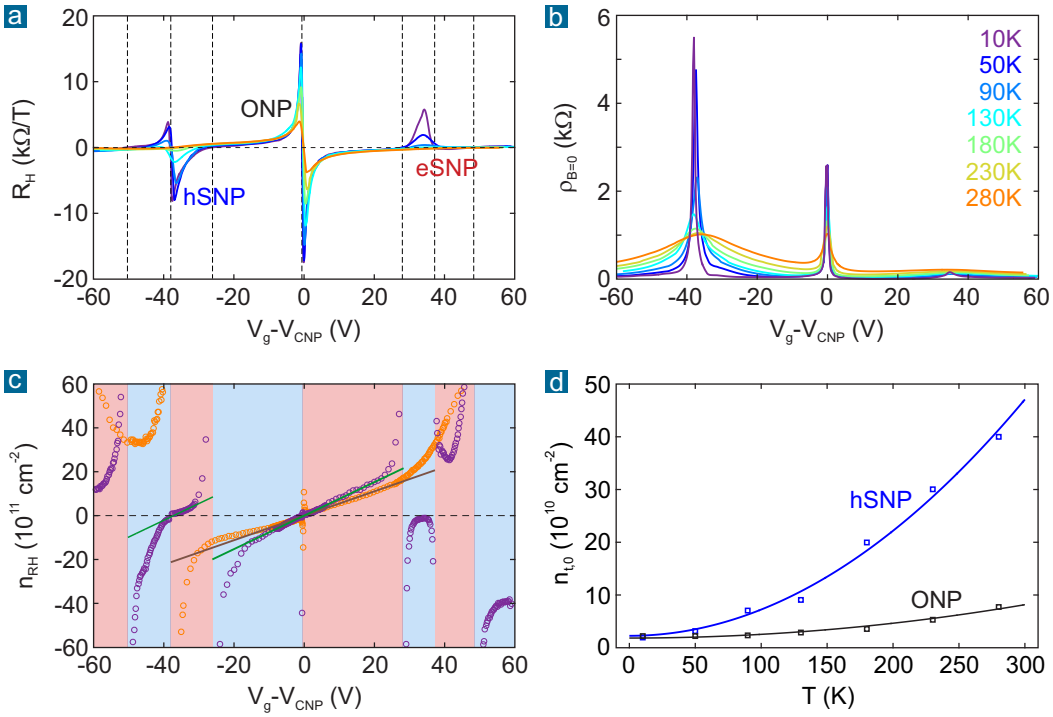

Figure 4.10. Effect of the Moiré superlattice on the weak field transport properties and its temperature dependence. Secondary neutrality points on the hole side (hSNP) and on the electron side (eSNP) emerge when the crystallographic planes of graphene and of (at least one of) the hBN layers are well aligned. The device presented here is encapsulated between two hBN flakes. The temperature dependence of the Hall coefficient and of the resistivity curves in absence of a magnetic field is shown in panel a and b, respectively. R_H was derived from fitting the Hall resistivity within ±50 mT magnetic field range. Both satellite neutrality points exhibit a strong temperature dependence and asymmetry. Panel c presents the carrier density at 10 K (purple) and at 280 K (orange) which we have derived from the measured Hall coefficient assuming the trivial single carrier model ($n_{RH} = 1/eR_H$). The transition between electrons and holes at 10 K is highlighted by the red and blue shaded background, respectively (the transitions are also marked in panel a by vertical dashed lines). The green lines correspond to the 10 K gate factor $\beta = 7.6 \cdot 10^{10}\,\mathrm{cm}^{-2}/\mathrm{V}$. The best fit to the n_{RH} curve at 280 K is marked by the brown line ($\beta = 5.6 \cdot 10^{10}\,\mathrm{cm}^{-2}/\mathrm{V}$). d: Temperature dependence of the residual total carrier density at the original neutrality point (ONP or CNP), determined by fitting the R_H curves, and at the hSNP, determined by the crossover point to the conductivity plateau in a double-logarithmic diagram.

linear. Hence, we can apply the same formula (Equation 2.24 in Section 2.2.2) to describe the thermal generation of charge carriers at the hSNP as in case of the ONP. We obtain $v_F^{hSNP} = 0.6 \cdot 10^6$ m/s for the Fermi velocity at the hSNP as a result of

the fitting. This reflects the fast thermal broadening which is already obvious when looking at the bare resistivity curves. On the contrary, the analysis of the ONP leads to $v_F^{ONP} = 1.6 \cdot 10^6\,m/s$ which agrees with other monolayer devices in absence of a Moiré superlattice. In case of the ONP, the residual carrier density values are derived from fits to the Hall coefficient curves[34]. Our result is in accordance with scanning tunneling spectroscopy experiments which find $v_F^{hSNP} = 0.5 \pm 0.1 \cdot 10^6\,m/s$ at the superlattice Dirac points [44]. Theoretical models predict Dirac cone replicas with halved group velocity (with respect to the main Dirac cone's group velocity) when graphene and the substrate material, e.g. hBN, exhibit a small lattice incommensurability [36]. In the supporting information of Ref. 45, the authors investigate the broadening of the resistivity peaks at the hSNP and at the ONP (for a graphene on hBN device). Based on their data we find $v_F^{hSNP} = 0.5 \cdot 10^6\,m/s$ and $v_F^{ONP} = 1.8 \cdot 10^6\,m/s$ when applying our fitting routine.

To conclude, the strong thermal broadening of the secondary neutrality points is governed by a linear band dispersion with halved group velocity (with respect to the main Dirac cone's group velocity). As a consequence, the transport features of the Moiré superlattice effect are drastically smeared out at room temperature and extend over a broad gate voltage range. Although the maximum Hall coefficients near the main CNP are still about $4000\,\Omega/T$ at $280\,K$, the measured Hall coefficients already deviate from the expected single carrier type picture at moderate gate voltages near the main CNP, e.g. $-10\,V$ at room temperature. This finding is more apparent when one compares the density values n_{RH} (orange data points) with the expected linear behavior governed by the gate properties (brown line in Figure 4.10c). The effect of the hSNP on the ONP at elevated temperature is also reflected by the resistivity curve which is clearly asymmetric. In addition, the temperature dependence of the main neutrality point's location $V_{CNP}(T)$ and of the gate factor $\beta(T)$ is found to be stronger in comparison to other stacks without Moiré superlattice. Although we do not yet understand the cause of the latter phenomenon, this finding and the other aforementioned issues lead to our recommendation that only non-aligned graphene-hBN heterostructures should be employed as magnetic field sensor. Besides the lower complexity of the system with respect to the theoretical model, the practical benefits are an undisturbed electron-hole symmetry and that the sensor can be tuned over a broad density range in controlled fashion.

34 Note, we also derived the $n_{t,0}$ values at the ONP from the double-logarithmic conductivity diagrams and obtained similar results.

4.5. Magneto-resistance at the CNP

We have shown in the first section of this chapter that the coexistence of electrons and holes around the CNP leads to a vanishing Hall coefficient which is fully understood within the framework of the two carrier type Drude model (which assumes a homogeneous system). This classical model further predicts the emergence of a strong positive magneto-resistance with quadratic magnetic field dependence under that condition. In Section 2.3.3 and Appendix B.2 we provide the full derivation of the magnetic field dependence of the longitudinal resistivity. For equal electron and hole mobilities $\mu_e = \mu_h = \mu$ and at the CNP $n_e = n_h$, it is given by

$$\rho_{xx}(B, \text{at CNP}) = \rho_{xx}(0)\,(1 + \mu^2 B^2)\,. \tag{4.9}$$

Here, $\rho_{xx}(0)$ is equivalent to the maximum resistivity at the CNP, $\rho_{CNP} = e\,n_{t,0}\,\mu_{CNP}$, which we have also introduced Section 2.3.3. The magneto-resistance at the CNP in the two carrier type Drude model then reads (repeating here Eq. 2.60)

$$MR \equiv \frac{\rho_{xx}(B) - \rho_{xx}(0)}{\rho_{xx}(0)} = \mu^2_{\text{two-carrier}}\,B^2\,. \tag{4.10}$$

The "two-carrier" subscript of the mobility indicates that we will use it as fitting parameter. However, we expect its value to be close to μ_{CNP}.

The measured magneto-resistance around the CNP for one representative hBN-encapsulated device is presented in Figure 4.11. First, we tune the system manually to the charge neutrality point by applying a voltage to the gate such that the resistivity reaches its maximum. Then, the magneto-resistance curve is recorded during a magnetic field sweep up to 200 mT. The as-measured data is shown in the inset of Figure 4.11a and typically exhibits a slight asymmetry regarding the polarity of the magnetic field. Following a common procedure in literature, we symmetrize the data and then calculate the magneto-resistance (MR) as defined in Equation 4.10. The amplitude of the magneto-resistance quickly drops when we tune the system away from the CNP (gray data points in Figure 4.11a) to larger effective gate voltages $V_g - V_{CNP}$ (orange and green data points). This behavior is in accordance with the classical Drude picture which tells us that there is no magneto-resistance in the single carrier type regime[35]. We have fitted the two carrier type Drude model (red lines) to the data using a common $\mu_{\text{two-carrier}}$ value for all gate voltages[36]. The

35 In the next section we will show that at lower temperatures other effects can cause the emergence of MR at higher densities. In the present case, at room temperature, the classical model holds.

36 Our procedure is as follows: We first fit the MR exactly at the CNP by using Equation 4.10. The fit parameter $\mu_{\text{two-carrier}}$ is then inserted in Equation 2.59 which describes the MR at any

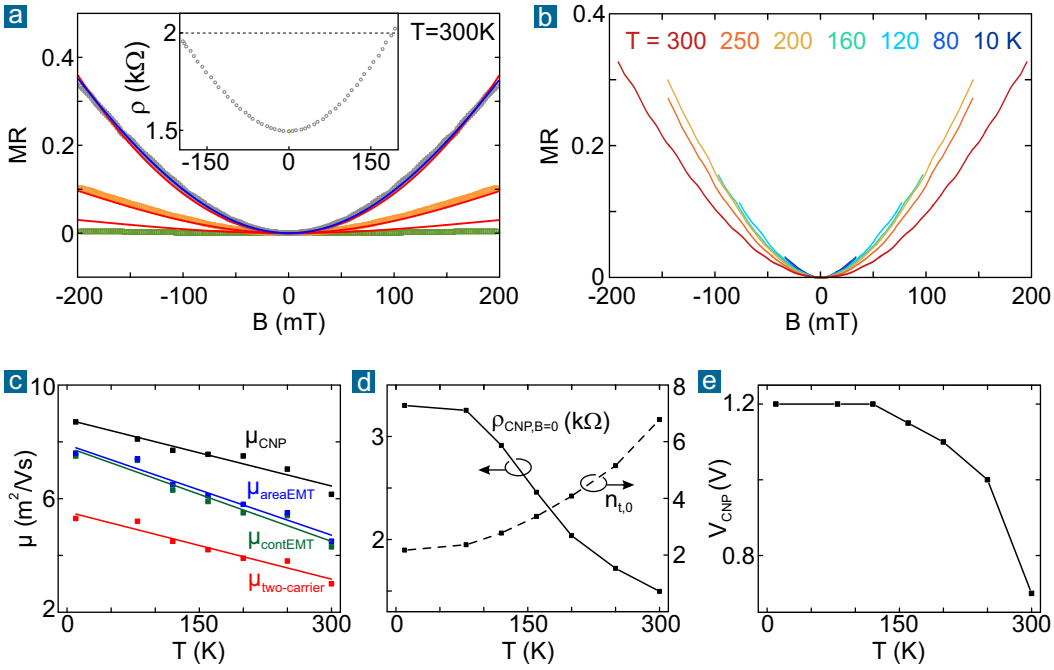

Figure 4.11. Observation of classical positive magneto-resistance at the CNP. All panels present one representative hBN-encapsulated device (with gate factor $\beta = 6.41 \cdot 10^{10}\mathrm{cm}^{-2}/\mathrm{V}$ and $n_{t,0} = 6.78 \cdot 10^{10}\mathrm{cm}^{-2}$ at 300 K). a: The inset shows the measured resistivity at the CNP against applied magnetic field. The data is symmetrized and then converted into the magneto-resistance MR by means of Equation 4.10. The main figure presents the MR at the CNP (gray), at an effective gate voltage of 1.5 V (orange), and at 3 V (green). The red lines are calculated employing the two carrier type Drude model where the mobility serves as fitting parameter (see main text). The blue line corresponds to the fit at the CNP by means of the area-fraction EMT formula. b: Temperature dependence of the MR at the CNP which allows us to deduce the dependence of the mobility used as fitting parameter (shown in panel c). d: Temperature dependence of the maximum resistivity at the CNP (in absence of a magnetic field) and the independently determined residual carrier density in units of $10^{10}\mathrm{cm}^{-2}$ (which allows us to calculate μ_{CNP}). e: Change of the charge neutrality point position with temperature.

quality of the fit is good at the CNP but overestimates the MR at higher densities. This might be related to the way we determine n_e and n_h (derived from $n_g(V_g)$ and $n_{t,0}$) and to the fact that the mobility also varies slightly with density.

combination of n_e and n_h. We know both density values to good approximation since we control $n_g = n_e - n_h$ by the gate and further, using the channel approximation, we can estimate that $n_t = n_e + n_h = \sqrt{n_{t,0}^2 + n_g^2}$. From separate Hall measurements we determine $n_{t,0}$ and the gate factor in order to calculate $n_g(V_g)$.

Earlier studies have reported that the two carrier type Drude model fails to capture the MR curve at the CNP for larger magnetic fields beyond one Tesla [158]. Furthermore, it is deficient in explaining the presence of weak but notable quadratic MR at higher densities for samples with strong disorder [110], that means with larger electron hole puddle amplitude (quantitatively described by the density fluctuation n_{rms}, see Section 2.2). The authors introduced effective medium theories (EMT) which account for the spatial inhomogeneity in graphene devices (while the two carrier type Drude model presumes the system to be homogeneous). Exactly at the CNP, for overall equal electron and hole concentrations, the magneto-resistance can be written in the following analytical forms (see Section 2.3.3 and Eq. 2.67)

$$MR_{areaEMT} = \sqrt{1 + \mu^2_{areaEMT} B^2} - 1\,; \quad MR_{contEMT} = A[n_g, n_{rms}]\, \mu^2_{contEMT}\, B^2 \tag{4.11}$$

The area-fraction EMT model accounts for spatial inhomogeneity by considering electron and hole regions which are assumed to be uniform inside [158]. Note that, for classically strong magnetic fields ($\mu^2 B^2 \gg 1$), the model predicts the magneto-resistance to depend linearly on the magnetic field. On the contrary, the continuous EMT model does not rely on this simplification but considers the continuous spatial distribution of the carrier density [110]. The additional pre-factor $A[n_g, n_{rms}]$ in the model depends in a complex manner on the gate controlled average carrier density and the disorder induced density fluctuations. Exactly at the CNP, the theory gives $A[0] = 0.5$ (see Section 2.3.3). Accordingly, both EMT models give very similar results at the CNP (for weak magnetic fields), while the continuous EMT model is found to capture better the observed weak MR at larger densities [110].

We have fitted both EMT models to our measured MR curves at the CNP. The blue curve in Figure 4.11a represents the fitting results of the area-fraction EMT model. Note, the continuous EMT model fit with $A = 0.5$ and fitting parameter $\mu_{contEMT}$ delivers exactly the same curve as the two carrier type Drude model (Equation 4.10), but the mobility value is by a factor of $\sqrt{2}$ larger than $\mu_{two-carrier}$. Within our measurement range, fulfilling reasonably the condition of weak magnetic fields, the MR curve remains quadratic and both the homogeneous two carrier type Drude and the inhomogeneous EMT models reproduce our data equally decently. This is in agreement with other studies [110, 158, 217, 357]. However, in contrast to the classical two carrier type Drude model which predicts a non-saturating quadratic behavior, several studies report that the MR at the CNP either saturates in case of low quality devices [158], or becomes linear in case of high quality devices [110, 217, 357], when the magnetic field range is extended to classically strong fields (when $\mu^2 B^2 \gg 1$). In particular, the latter case of strong linear magneto-resistance (LMR) has been observed between 150 K and He base temperature [217, 357]. The

mechanism behind the appearance of a linear magneto-resistance in graphene is still under debate[37]. In this work, we have restricted us to weak magnetic fields focusing mainly on the behavior at room temperature. Thus, we observe just an onset of LMR in case of very high mobility hBN-encapsulated devices and at lower temperatures. Due to the increase in mobility, we approach the regime of classically strong magnetic fields already within the measurement range of $\pm 200\,\mathrm{mT}$. However, the effect is accompanied by other features such as universal conductance fluctuations at low temperature (not shown). Hence, we focus on the analysis of the quadratic MR in the following paragraph and suggest further experiments to reveal unambiguously the origin of the LMR[38].

In Figure 4.11b we show the temperature dependence of the quadratic part of the observed magneto-resistance at the CNP[39]. For each temperature, the measured curves are fitted with all three MR models discussed earlier. The resulting mobility values are presented in Figure 4.11c together with the temperature dependence of μ_{CNP}. The latter is obtained from the measured maximum resistivity in the absence of a magnetic field and the $n_{t,0}$ values which we have determined from the Hall coefficient curve fitting. In general, we find the temperature dependence to be linear for all mobility definitions but the absolute values vary. In particular, $\mu_{\mathrm{two-carrier}}$ and μ_{CNP} exhibit the same slope ($-0.008\,\mathrm{m^2\,/\,Vs\,K}$) but the absolute values differ unexpectedly by a factor of two.

37 The linear magneto-resistance effect quickly disappears when the carrier density is tuned away from the CNP which excludes the classical random resistor network model as possible explanation [216, 217, 358]. As already discussed, the (area-fraction) effective medium model predicts a linear magnetic field dependence for classically strong fields [158, 199]. This would be supported by the existence of the electron hole puddle landscape which in particular becomes dominant at low temperature. Another recently proposed model universally describes the emergence of LMR in two-component systems with equal concentrations of electrons and holes by recombination effects in confined sample geometries [216, 217, 357, 359]: Exactly at the CNP, the compensated Hall effect is accompanied by electron-hole recombination at the sample edges. When the characteristic recombination length becomes similar to the channel width W, electrical transport is primarily dominated by classical edge currents which is shown to lead to LMR [216, 359]. The condition of narrow channels (with a width of only few micrometer) is typically fulfilled in hBN-encapsulated samples since the device dimensions are often limited by bubbles, as in the case of Refs. 217, 357

38 The improved stacking procedure now allows producing high quality devices whose dimensions are not anymore restricted to small channel widths due to bubbles. Hence, we suggest to examine a series of devices with substantially varying channel widths (at low temperature and with increased magnetic field range such that the LMR is more pronounced). If the recombination picture is correct, the LMR is expected to occur only for narrow devices. But if the LMR is observed in wide devices (and only at low temperature when the spatial inhomogeneity is relevant), the effective medium theory is a possible candidate.

39 For each temperature, we have limited the data range which was considered for the quadratic fitting in order to exclude the transition to the linear dependence and the emergence of other effects, especially at the two lowest temperatures.

We finish this section by commenting on the question how (im-)practical it is to exploit the magneto-resistance at the CNP in graphene for magnetic field sensing applications at room temperature. First, for small magnetic fields, the response is not linear but quadratic. This implies that the expected signal level is low compared to the noise level (which reaches a maximum at the CNP, as discussed in Chapter 5). In addition, the mobility changes with temperature and so the quadratic pre-factor. Second, the system has to be tuned to and kept stable at the CNP for the magneto-resistance to arise. As shown in Figure 4.11e, the position of the CNP, in terms of the necessary voltage which has to be applied, changes considerably with temperature. Third, the resistivity value at the CNP in the absence of a magnetic field is strongly temperature dependent (in a complex manner due to thermal activation of charge carriers and $\mu_{\mathrm{CNP}}(\mathrm{T})$). This implies that the zero magnetic field reference would have to be re-calibrated continuously. In sum, we believe that these issues preclude reliable sensing of small magnetic fields based on graphene's intrinsic[40] magneto-resistance response. Hence, this work focuses on graphene based Hall sensors instead as they offer a broader range of possible working points (as discussed in Chapter 5).

4.6. Comment on transport phenomena beyond the diffusive picture

In the previous sections of this chapter we have disregarded all transport phenomena beyond the classical two carrier type Drude model which was found to describe the diffusive regime properly. Hence, for the main scope of this work, the application of graphene for magnetic field sensing at room temperature, the model is sufficient. Recent reports in the literature, however, have provided strong evidence that mesoscopic transport phenomena emerge in high quality, e.g. hBN-encapsulated, graphene devices. While ballistic effects become dominant at low temperatures, viscous transport phenomena can play an important role at moderate temperatures (however, typically well below room temperature). The theoretical framework of these transport regimes and the requirements for their emergence are discussed in Sections 2.4 and 2.5, while in the following we focus on the key experiments.

40 We emphasize that graphene still might be employed as magnetic field sensor to detect large magnetic fields (of several Tesla) when observing the extraordinary magneto-resistance effect in special device geometries [360, 361]. Moreover, a large non-local magneto-resistance has been reported in few-layer graphene on hBN heterostructures [362]. Then, large magnetic fields are found to induce the Ettingshausen-Nernst effect in these structures. However, for the purpose of detecting weak magnetic fields we rely on the intrinsic MR response of graphene which has been shown here to be impractical.

In the absence of a magnetic field, experimental evidence of ballistic transport at low temperature involves the measurement of a negative bend resistance in cross- or square-shaped devices [10, 195, 203, 218] or the observation that the conductivity is limited by the channel width and exhibits a square root dependence on the carrier density at higher densities (see Section 3.2.2 and Refs. 68, 205). Ballistic features in magneto-transport were further identified by Masubuchi and coworkers [162]. The authors find the magneto-resistance curves obtained on hBN-supported, wire-like graphene devices (width $W \approx 1\mu m$ and length $L \approx 3\mu m$, measured in two-terminal configuration) to possess a characteristic peak structure at low temperature and high density when the magnetic field is swept up to 250 mT (see Figure 2.8a in Section 2.4). The characteristic magneto-resistance peak is then explained by the interplay between the varying cyclotron radius (when the magnetic field is swept) and the scattering at the channel boundaries which constitutes the dominant scattering mechanism in the ballistic regime. Other experiments on extremely clean (hBN-encapsulated) graphene devices have found clear indications of a fluid-like behavior of the charge carriers at moderate temperatures (typically between 70 and 200 K). The viscous transport regime reveals itself in negative local resistances (in vicinity of the current-injecting contact due to whirlpool-like electron backflow) [210], in superballistic transport along constrictions [205], or in the breakdown of the Wiedemann-Franz law [79]. All these experiments were performed in the absence of a magnetic field. Recently, during the analysis of our measurements and preparation of this manuscript, several theoretical studies were published proposing observables for magneto-hydrodynamics which so far could not be studied experimentally due to the lack of suitable systems [209, 214, 215, 363, 364]. Possible observables are anomalous features in the Hall resistivity (related to the Hall viscosity) and the emergence of a negative magneto-resistance in the single carrier type regime. In particular, Scaffidi and coworkers (Ref. 209) have calculated the magneto-resistance curves and the Hall resistivity curves both for the ballistic and for the hydrodynamic case (see Figure 2.8 in Section 2.4).

Since our best hBN-encapsulated devices possess comparable quality and dimensions to the samples used in literature, one may raise the question whether our measurements below room temperature are affected by these phenomena beyond the diffusive picture. Although our experiments were not explicitly designed for this purpose[41], we briefly present our observations. As discussed in Section 3.2.2

41 First, we chose a narrow magnetic field range with a maximum of $100 - 200\,mT$ since we already have restricted our analysis of the Hall coefficient to $\pm 50\,mT$ to avoid the (classical) non-linearities in the Hall resistivity curves when $\mu^2 B^2 \gg 1$ and electrons and holes coexist around the CNP (see Figure 4.1e). Second, we used conventional Hall bar (or cross) geometries (with $L/W > 3$ and a width of some micrometer). The width of the voltage probes was on the order of one micrometer and the probes were not closely spaced (with a distance of at least one

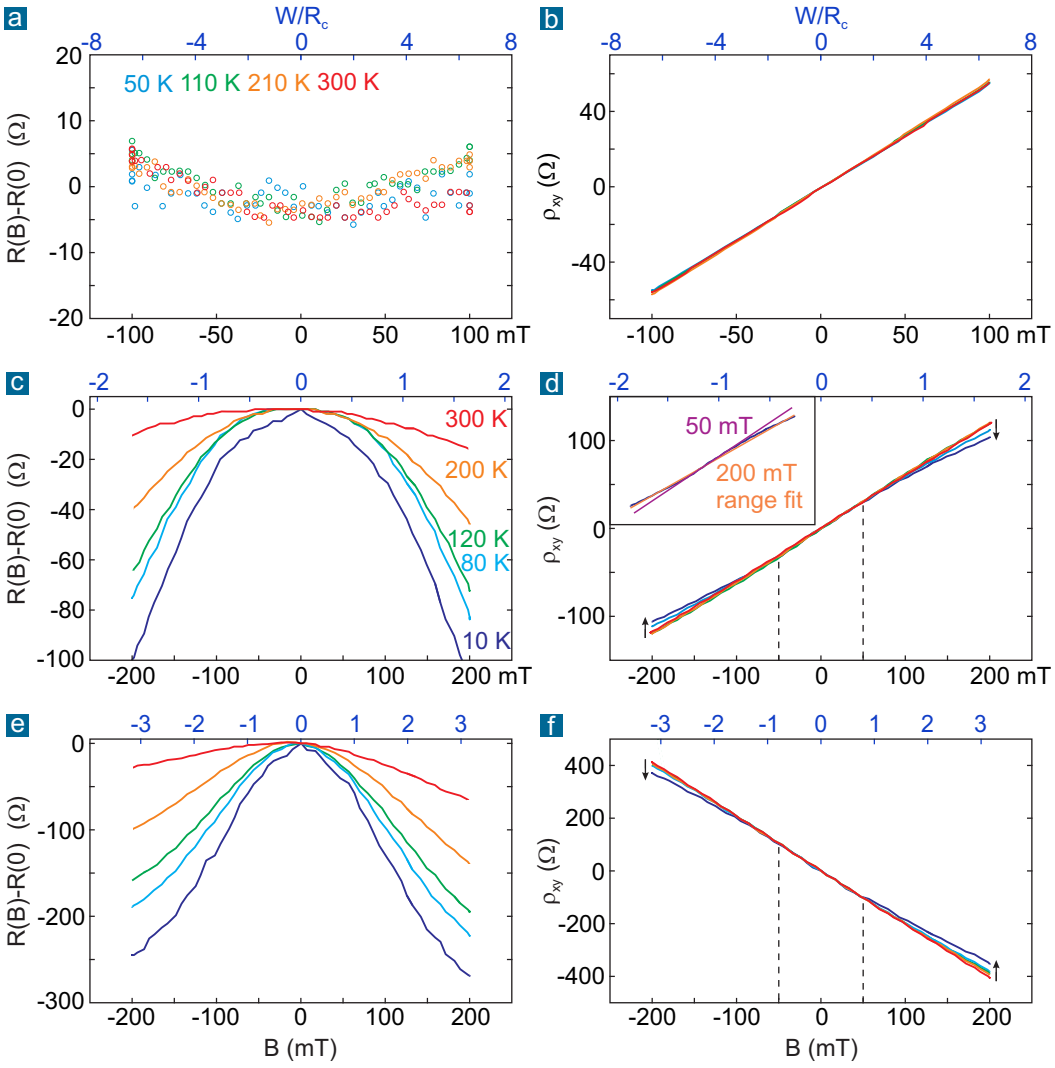

Figure 4.12. Emergence of negative magneto-resistance and non-linearity in the Hall response in the single carrier type regime, not captured by the classical diffusive picture. We present here as-measured data obtained from a bare graphene on SiO_2 device with channel width $W = 8\,\mu m$ and hole density $n_h = 1 \cdot 10^{12}\,cm^{-2}$ (panels a , b) and from a hBN-encapsulated device with $W = 1\,\mu m$ and $n_h = 1 \cdot 10^{12}\,cm^{-2}$ (panels c, d) or $n_e = 3 \cdot 10^{11}\,cm^{-2}$ (panels e, f). Channel length L is $24\,\mu m$ and $7\,\mu m$, respectively.

micrometer). We also employed the conventional contact configuration for magneto-transport measurements. The current was applied along the main channel of the Hall bar while the side contacts were used to measure the longitudinal and the Hall voltage. Hence, we did not measure the (negative) bend resistance as direct proof of ballistic transport at low temperature, except

and shown in Figure 3.8, the measured conductivity at low temperature and high density typically approaches the ballistic limit. However, we do not observe the predicted characteristic magneto-resistance peak which we believe is due to our different device design (see footnote). Instead, we find the magneto-resistance curves in some cases to be rather complex, likely due to the interplay of several effects, e.g. mesoscopic conductance fluctuations, and due to the influence of the Hall voltage on the measured longitudinal voltage. Nevertheless, the common trend is that a negative magneto-resistance develops at higher densities away from the CNP (in the single carrier type regime). Its amplitude increases with decreasing temperature and as the carrier density is lowered (up to the point where the system enters the two carrier regime and strong positive classical magneto-resistance arises). This observation is also accompanied by the onset of sublinearity in the Hall resistivity curves which manifests for magnetic fields above 50 mT. Both phenomena cannot be explained within the classical model because then no magneto-resistance is expected and the Hall response is perfectly linear in the single carrier type regime. Figure 4.12 shows the results of two different samples representing the clearest example of the described behavior. Both phenomena emerge only in hBN-encapsulated graphene with high mobility and narrow channel width (limited due to manufacturing reasons). They are absent in low quality, bare graphene on SiO_2 devices for which the classical diffusive picture holds.

In order to compare our measurements to the aforementioned theoretical predictions of Scaffidi and coworkers in Ref. 209 (shown in Figure 2.8 in Section 2.4), we have added a second x-axis with the dimensionless parameter W/R_c in all panels of Figure 4.12. It is determined by the ratio between the channel width W and the cyclotron radius $R_c = \hbar\,(e\,B)^{-1}\,\sqrt{\pi n}$. In our measurements we vary the value of W/R_c between zero and (up to) three by sweeping the magnetic field. In the theory of Ref. 209, the bulk case is defined as the ohmic (diffusive) case, when momentum-relaxing scattering dominates ($L_{MR} \ll W$). This corresponds to our measurements at room temperature. Boundary scattering in the ballistic regime prevails in transport (when W is the smallest length scale), while the hydrodynamic case is approached when the momentum-conserving carrier-carrier scattering length is the smallest length scale

for one square-shaped device which indeed showed a negative bend resistance (not shown). The measurement of the negative local resistance as evidence of the fluid-like behavior requires closely spaced voltage probes, a narrow current injector, and an unconventional current flow through the sample [210, 364]. Accordingly, we also do not observe the latter in our measurement configuration. Moreover, our Hall bar geometry with side contacts and the four-terminal measurement of the resistance likely obscure the observation of the boundary scattering related magneto-resistance peak in the ballistic regime since Masubuchi et al. (Ref. 162) used rather short samples and measured the magneto-resistance in two terminal configuration, without any lateral (Hall) contacts.

of the system (see also Section 2.4). If at all, the latter condition is only fulfilled at intermediate temperatures (compare to Equation 2.15 in Section 2.1.7).

Based on our measurements, we can exclude that we reach the purely ballistic or hydrodynamic limit in our system (which was simulated in the theory as ideal cases). The theoretical study in Ref. 209 predicts that the Hall resistivity becomes smaller than for the bulk (diffusive) case when W/R_c approaches zero. We observe the exact opposite. The Hall resistivity slope remains linear and the same at all temperatures for small magnetic fields (< 50mT), but then becomes sublinear for larger fields. This sublinearity is particularly pronounced at low temperature. Although the hydrodynamic regime is characterized by negative magneto-resistance, too, qualitatively similar to our data, it does not explain why we see a steady increase of the amplitude when the temperature is lowered down to 10 K. At this low temperature, the carrier carrier scattering length is larger than the system dimensions and hydrodynamic effects cannot be invoked to explain the negative magneto-resistance and the sublinearities in the Hall response for larger magnetic fields. Moreover, as mentioned above, we also do not observe the characteristic peak in the magneto-resistance trace which would indicate the purely ballistic case. Hence, the origin of the observed behavior remains unclear[42]. Additional experiments, beyond the scope of this work, would be required to gain a complete understanding of the matter[43].

Note. After finishing this chapter of the manuscript, a study was published claiming the measurement of the Hall viscosity of graphene's electron fluid [364]. The authors found the viscous fluid to respond to weak magnetic fields (< 40 mT) by generating an electric field opposite to that induced by the classical Hall effect. The suppression of the normal Hall response has its maximum around 100 K but quickly decays with increasing temperature and vanishes around room temperature. Moreover, the effect is only observable near to the current injection probe which itself is required to be narrow (having a sub-micrometer width). As also emphasized by the authors, these

42 We further can exclude quantum interference effects since they are dominant within a much smaller magnetic field range (see Appendix B.1). The sublinearities arise when $W/R_c > 1$ which means that the cyclotron radius becomes smaller than the channel width. Although we do not yet detect Shubnikov de Haas oscillations, the sublinearities might be also related to the onset of Landau quantization. Other possible explanations for the negative magneto-resistance involve electron electron interaction in the diffusive regime [76] or other quasi-classical theories in two dimensions as discussed in Refs. 365, 366.

43 The sample design has to be optimized for magneto-resistance measurements or for local voltage measurements. The new stacking method allows obtaining larger Hall bars without being limited by bubbles. Furthermore, the magnetic field range should be increased in order to access a larger range of the dimensionless parameter W/R_c (which can be easily done but was not necessary for our primary purpose).

requirements make it difficult to observe the Hall viscosity in conventional geometries, or in other words, hydrodynamic effects are not expected to affect our Hall voltage measurements. However, the study also reports the emergence of negative magneto-resistance which is attributed to viscous transport above 200 K and to the appearance of skipping orbits in a magnetic field suppressing backscattering at low temperature [364].

To conclude, the here briefly discussed phenomena beyond the diffusive picture are highly interesting for basic research since graphene constitutes an excellent system for studying magneto-hydrodynamics. However, we can now safely say that for the main scope of this work the classical diffusive picture holds since we focus on the room temperature magneto-transport of graphene in order to employ it in future magnetic field sensing applications. Moreover, as we have shown, the Hall resistivity traces at low temperature and at high densities remain linear and identical[44] to the diffusive case for small enough magnetic fields, e.g. below 50 mT. Otherwise, when a larger magnetic field range is considered to determine the Hall coefficient by means of a linear fitting model, the obtained Hall coefficients at low temperature are smaller than when considering only fields up to 50 mT (as shown the inset of Figure 4.12d). Note that we already have restricted us to this magnetic field range when determining the Hall coefficients in Sections 4.1 and 4.3 in order to avoid that the Hall coefficient depends on the magnetic field (as shown earlier, within the classical two carrier type Drude model, this can occur in the two carrier regime when $\mu^2 B^2 \gg 1$). Hence, for our conventional Hall bar devices, our procedure to obtain the Hall coefficients through linear fitting of the Hall resistivity traces remains valid also at high densities and over the whole temperature range, independent from the (mixed) transport regime. Accordingly, we can also exclude ballistic or hydrodynamic effects as cause of the observed asymmetries in the Hall coefficient values between the electron and hole side at high densities (see Section 4.2). Even if relevant, these effects would be identical for both electrons and holes.

44 Here, "linear" refers to that we do not observe any non-linearity in the Hall resistivity traces when W/R_c (the magnetic field) is tuned to zero, in contrast to the theoretical predictions of the purely ballistic and hydrodynamic case (compare to Figure 2.8). Identical traces, e.g. slopes, are only obtained when one compares the Hall traces for fixed (high) carrier density. This implies that one has to account for the slight temperature dependence of the charge neutrality point V_{CNP} and the gate factor β (due to extrinsic effects). Accordingly, the slope of the Hall resistivity trace for fixed gate voltage certainly changes with temperature (but not for fixed, high enough density).

Chapter 5

Technical assessment of graphene based Hall elements

As shown in the previous chapters, graphene can be tuned easily to small carrier densities through electrostatic gating before eventually the two carrier regime is reached around the CNP where electrons and holes coexist. When the extrinsic disorder level is low, i.e. in high quality, hBN-encapsulated graphene, the residual total carrier density at the CNP at room temperature is mainly governed by thermal activation of carriers and maximum Hall coefficients above 5000 V/AT can be obtained. This means that graphene outperforms all other materials employed in Hall elements, i.e. Si and III-V compound semiconductors, in terms of maximum magnetic field sensitivity. Moreover, room temperature mobility values in such high quality graphene reach record values and resistivity values remain moderately low over the entire charge carrier density range, peaking at around 1 kΩ at the CNP.

These properties make graphene, at first glance, a promising candidate to substitute today's state of the art transducer materials in commercial Hall sensors. The latter are widely used in applications requiring magnetic field measurement in the range between some μT, i.e. geomagnetic field sensing, and some mT (or, in fewer cases also some ten or hundred mT). In this chapter, we will investigate in detail the performance of graphene based Hall transducers around room temperature, assessing all important technical parameters for this prevalent type of magnetic field sensor.

We will show that both the already discussed sensitivity, i.e. the Hall coefficient, as well as the Hall voltage noise determine the minimum magnetic field which can be detected given a certain operating current (or voltage). The latter property is also referred to as magnetic field resolution or detectivity. It constitutes the most important performance indicator for many applications, followed by the sensitivity, the Hall voltage offset at zero magnetic field and the linearity. Thermal and 1/f

noise are typically the main contributors to the overall electrical noise level in Hall transducers. They also prevail in graphene based devices as will be discussed in the following sections. Since 1/f noise scales inversely with the total number of charge carriers and with the frequency, the attainable magnetic field resolution for a given bias current depends on the operating frequency, carrier density, and volume or area (for 2D materials) of the Hall element. The 1/f noise also increases with the applied current. Hence, the magnetic field resolution cannot be improved any further by simply increasing the current when 1/f noise dominates. Then, any gain in the absolute Hall voltage signal (Hall coefficient times current) is canceled out by an equivalent increase in the Hall voltage noise amplitude, and the signal to noise ratio remains the same. All these (inter)dependencies will be studied in detail in the course of this chapter.

While typical sensitivity (Hall coefficient) values can be easily compared among different Hall transducer materials, a comparison in terms of magnetic field resolution is more difficult since the latter depends on the exact operating conditions and device dimensions. The operating parameter space until 1/f noise dominates is typically extended in commercial devices through choosing a large enough device area and a high enough operating frequency. Then, the magnetic field resolution can be improved by increasing the current, and so the absolute sensitivity, as long as the 1/f noise remains below the thermal noise floor (assuming that current heating is negligible). However, this approach often encounters practical limits such as strict requirements on power consumption or maximum supply voltage when battery powered, available footprint on the chip, and operating frequency (increasing complexity and power consumption of the other integrated electronic components aside from the Hall transducer).

In order to be able to study the properties of graphene having the highest quality possible, that means determining truly its intrinsic limits, we rely on exfoliated graphene and hBN flakes for fabrication and encapsulation of the Hall elements. This limits, however, the dimensions of the Hall transducers under test. They measure at most few tens of micrometer, in contrast to the typically much larger Hall elements made from other materials that are reported elsewhere in the literature and used in commercial Hall sensors. Therefore, we will also provide noise and magnetic field resolution data that is normalized regarding device geometry and absolute dimensions. Using the Hooge model for 1/f noise in graphene allows us to extrapolate to typical device dimensions used in today's commercial Hall elements[1]. Furthermore, while the doping is mostly fixed in case of other materials through dopants, e.g. in Si

1 Of course, the future realization of such large area, high quality graphene based Hall transducers relies on progress in large-scale fabrication methods of graphene and, in particular, its

Hall transducers, the carrier density in graphene Hall elements can be easily tuned through electrostatic gating. This enlarges the parameter space we can and need to explore: The (normalized) Hall voltage noise and the magnetic field resolution will be studied as a function of frequency, current and density. This rich data set allows us to compare our results with values reported elsewhere on both graphene as well as other materials, answering the question whether graphene performs better under similar operating conditions. Moreover, it can be estimated which ultimate benchmarks could be reached by large-area, high quality graphene Hall elements, if all operating parameters and conditions were optimized.

This chapter starts with a general overview on magnetic field sensing. We briefly introduce several sensor types, whose working principle are based on different physical effects, and categorize them by their typical fields of application. When discussing magneto-resistive sensors, we also comment on room-temperature magnetoresistance effects in graphene which, in principle, constitute an alternative to the Hall effect for measuring magnetic fields with graphene based devices (see also Section 4.5). The second section gives a general introduction to Hall effect based magnetic field sensors and their key performance indicators, in particular the magnetic field resolution. Modern Hall sensors are sophisticated ASICs and do not only consist of the Hall transducer or element, i.e. the actual sensing part. The Hall element, or sometimes even several connected Hall elements, is combined with additional electronic components in one chip, for example for driving the Hall element, readout, A/D conversion, and signal processing. Such integrated Hall sensor devices offer many possibilities to improve the performance of the Hall transducer material. These include offset and noise reduction techniques, linearization and compensation of thermal drifts. In this context, we emphasize that the data presented in the subsequent graphene specific sections were always acquired on "bare" graphene Hall elements without any further optimization through on-chip electronics. Thus, it should be compared only to other materials for which characteristics are known on the Hall transducer level. The on-chip integration of graphene Hall elements and the layout of application specific ICs goes beyond the scope of this work.

In the third section, we discuss the theory of common types of noise. The empirical Hooge model for 1/f noise is introduced for graphene Hall transducers which are governed by an interplay between thermal and 1/f noise. Combining this specific noise model and the absolute sensitivity, a formula for the magnetic field resolution

ideal substrate hBN. Therefore, at the end of this chapter, we will briefly review recent developments in large-scale synthesis of 2D materials as well as proposed ways to integrate these unconventional materials into established semiconductor fabrication processes.

can be given which will be verified by our measurement data presented in the subsequent sections. Moreover, we review the most important findings about 1/f noise in graphene in the literature, at which noise voltage measurements in longitudinal, i.e. resistance fluctuations between source an drain, and transverse configuration, i.e. between the Hall contacts, will be distinguished. In the fourth section, we describe our custom-built setup to study all important properties of graphene Hall elements and specify its main properties. It consists of carefully selected measurement equipment and a self-designed, properly shielded PCB sample holder equipped with integrated low noise pre-amplifiers. It offers a very low instrumental noise floor and the possibility to control the magnetic field precisely through a built-in reference sensor as well as to vary the temperature between room temperature and about 80°C.

The following sections 5 and 6 then discuss our measurements of the Hall voltage noise in graphene Hall transducers and the obtained magnetic field resolution, respectively. The interplay between 1/f and thermal noise will be shown. All above mentioned (inter)dependencies of the operating parameters and the applicability of the Hooge model will be demonstrated. As in the previous chapter, bare graphene on SiO_2 and high-quality, hBN-encapsulated graphene devices as well as monolayer and bilayer devices are compared. Finally, the temperature dependence is investigated. Section 7 then summarizes all technical specifications of graphene Hall transducers we studied in this work. We will identify four key material parameters and a set of in this work validated models, i.e. the two carrier Drude, the channel approximation and the Hooge model, that allow us to simulate graphene Hall transducers with larger area, comparable to the footprint of today's commercial Hall elements, and assess the potential performance limits under optimized operating conditions. Both our experimental results as well as the outcome of these simulations will be compared to other studies on graphene Hall elements as well as other material systems. Finally, section 8 discusses possible routes to realize such high quality, large-area graphene devices in the near future and integrate them into existing semiconductor fabrication processes.

5.1. Magnetic field sensing in commercial applications

Magnetic field sensing can be conducted in versatile ways based on a multitude of physical effects. In general, magnetic field sensors or magnetometers can be classified into two main categories. Scalar magnetometers measure the absolute value of the magnetic field but not its direction, while vector magnetometers detect a particular spatial component of the magnetic field, relative to the orientation of the magnetometer's active sensing part. For example, Hall elements are sensitive to the out-of-plane magnetic field component, while anisotropic magneto-resistance (AMR) devices measure the in-plane component. In modern commercial devices, e.g. in an electronic compass, several (types of) vector magnetometers can be also arranged in one package to allow for magnetic field measurement along all three orthogonal directions. These 3-axis magnetometers are often combined with a multi-axis accelerometer unit for tilt compensation, a multi-axis gyroscope unit and a micro-controller unit for on-chip processing of the raw data streams (referred to as integrated sensor fusion or system in package).

Magnetometers that are based on different physical working principles may differ in terms of absolute sensitivity, dynamic magnetic field range, magnetic field resolution, physical size, cost, power consumption, complexity in terms of production and operation (e.g. cryogenic temperature needed), longevity, linearity error, temperature range and drift, read-out principle (optically or electronically), read-out speed (e.g. possible output data rate), and other parameters. Therefore, the optimal choice of a magnetic field sensor depends on the list of requirements given by the specific application. A universal device for all uses, having a very high resolution and dynamic range at ultra-low power consumption while being inexpensive, operable over a wide temperature range, and small in physical size, does not exist. A broad and up-to-date overview of all relevant types of magnetometers and their typical uses is given in the book of Grosz and coworkers [367]. It covers search coil, fluxgate, giant magneto-impedance, SQUID, magneto-optical, optically pumped, as well as magnetic resonance based magnetometers (and more). Moreover, new approaches such as magnetometry with nitrogen-vacancy centers in diamond are reviewed. These magnetometers will not be discussed further in this work since their physical size and cost, if they are commercially available at all, are orders of magnitude larger than feasible for most industrial or consumer electronics applications. However, they allow for measuring magnetic fields (far) below one μT and possess ultra low noise levels (i.e. an excellent magnetic field resolution). Fields of application for such high-end magnetometers include, among others, geophysical and astronomical research, archeology, aerospace and defense, as well as healthcare (e.g. in medical imaging).

For the vast majority of industrial or consumer applications, low-cost vector magnetometers of small physical size and with a magnetic field range between some μT and some (tens of) mT are required. In case they are battery powered, the maximum supply voltage is limited and the power or, to be more precise, the current consumption has to be minimized. The upper industry standard corresponds to a magnetic field sensor having a price tag on the order of one US Dollar and a physical size on the order of few cubic millimeter that is capable of measuring magnetic fields in the range of $\pm 1,000\,\mu$T with a resolution better than one μT (quite often even $\leq 300\,$nT) and small thermal drift between $-40°$C and $85°$C. These upper standard specifications are in particular relevant for electronic compass applications involving geomagnetic sensing (earth's magnetic field $\approx 60\,\mu$T). Use cases are navigation and orientation in mobiles, wearables, gaming devices, drones etc. (that means primarily in consumer electronics). These devices are typically offered as 3-axis magnetometer in one small package with on-chip processing and digital output[2]. Then, the output data rate (ODR) is another important parameter and ranges between $10 - 160\,$Hz. The current consumption, taking into account both the sensing element and the other on-chip components, is on the order of few hundred μA at a supply voltage of up to $3.6\,$V (when Li ion battery powered). In conclusion, the main figures of merit of these digital electronic compass ICs are the magnetic field resolution, current consumption and ODR.

Aside from the compass application in consumer electronics, magnetic field sensors are also used in many other applications. Functionalities offered by the sensors are switches or latches, current sensing, (rotation) speed measurement and angular or linear position sensing in household and industrial machines, robotics, and automotive. Examples are anti-lock breaking, joysticks, washing machines, process control, motor control, and many more. Due to this broad spectrum of applications the specifications of the commercial devices vary much stronger and the requirements in terms of magnetic field resolution, physical size and power consumption are typically less demanding than for the high-end electronic compass. Since these magnetometers are often used in combination with a permanent magnet (e.g. to detect movements or rotations), the to be measured magnetic fields are larger than for geomagnetic sensors ($\approx \pm 10\,$mT range). Higher supply voltages, e.g. 24 or 48 V, and both digital as well as analog output devices are common. One single-axis magnetometer alone is often sufficient or, if necessary, 3-axis sensors are simply made by combining several comparatively bulk PCB pieces each containing one single-axis

2 Examples for such 3-axis digital compass devices are HMC5883L from Honeywell (AMR sensor), MMC3416 from Memsic (AMR sensor), YAS539 from Yamaha (GMR sensor), BMM150 from Bosch (Hall sensor in z, proprietary Flipcore technology for x and y), LIS3MDL (TMR sensor) or LSM303AGR (AMR sensor) from ST Micro, Xtrinsic MAG3110 from NXP (TMR sensor), or AK09918 from AKM (3-axis Hall sensor).

device (since extreme miniaturization is not required)[3]. However, as a consequence of the lower level of performance requirements, the cost is more decisive.

Si CMOS integrated magnetic field sensors

In order to meet the particularly challenging requirements in terms of ultra low cost and relatively small physical size imposed by all large-volume market segments, magnetometers have to be producible in very large quantities and with high yield regarding device-to-device performance stability. This can be realized only through semiconductor batch fabrication processes. Moreover, as already mentioned, state-of-the-art devices are equipped with several additional integrated electronic components that optimize the performance of the bare sensing element. These components are fabricated in Si CMOS technology. Therefore, either the fabrication of the sensing element can be integrated easily into the standard Si CMOS front-end-of-line (FEOL) steps, e.g. in case of Si Hall elements, or it has to be at least compatible, that means not degrading the wafer properties, when the fabrication happens during the back-end-of-line (BEOL) steps. This poses strict constraints on the maximum temperature and the choice of chemicals or materials, e.g. due to contamination issues, that can be used in the deposition of non-Si materials for the sensing element. This applies to both direct growth as well as transfer or wafer bonding processes.

Technological progress happens gradually with better performing integrated magnetic field sensors being released every few years. The main drivers for improvement are new materials beyond Si in Hall effect devices, requiring the integration of III-V compound semiconductors such as GaAs and InSb, or the exploitation of different types of magneto-resistive effects in XMR sensors, requiring the integration of ferromagnetic materials in Si CMOS. Typically, this is realized through multi-wafer stacking during BEOL, or the sensing element is deposited on top of the Si CMOS chip through low temperature processes (e.g. sputtering). As will be discussed briefly in the last section of this chapter (Section 5.8), integrated graphene Hall elements are expected to be fabricated in a similar manner and during BEOL, i.e. through transferring CVD grown graphene onto Si CMOS wafers with already integrated components.

Hall effect devices, and in particular Si based ones, still have the biggest share in the market due to competitive cost and high level of integration. However, the

3 Examples are SS39ET from Honeywell (Hall sensor), A1324 from Allegro (Hall sensor), CSA-1V from Sentron (Hall sensor), MLX91205 from Melexis (Hall sensor), MFS-3A from Ametes (3-axis Hall sensor PCB stack), STJ-3D from Micro Magnetics (3-axis TMR sensor PCB stack)

historically leading Si Hall technology has reached its physical limits. Magneto-resistive sensors (XMR) perform better in terms of magnetic field resolution at lower power consumption and exhibit the highest market growth rates (AMR, GMR and recent TMR technology will be described briefly in the following subsection). Hence, the more demanding the requirements of a certain use case, the more unlikely conventional Si Hall sensors are employed. GaAs or InSb based Hall sensors that also show better performance than Si Hall devices are mainly used in specific niche applications because of higher cost. Therefore, an investigation of the performance of graphene based Hall elements, as conducted in this chapter, is of great importance. It answers the question whether Hall effect based magnetic field sensors can remain competitive on the long term in use cases with challenging requirements. Of course, the final answer to this question depends also on cost competitiveness, that means the scalability of the fabrication process of integrated graphene based Hall sensors, which is beyond the scope of this work.

As an alternative to the market dominating Hall or XMR devices, some manufacturers also offer MEMS based magnetometers which sense either electronically or optically the mechanical motion of a micro-structure due to the Lorentz force that acts on a conductor carrying a current in the presence of a magnetic field. Moreover, MEMS fabrication technology allows to produce a miniaturized transformer. Then, the timing of the voltage spike in the secondary coil can be related to the externally applied field when the current in the primary coil, containing a magnetic material, is reversed periodically (proprietary, so-called Flipcore technology developed by Bosch Sensortec). MEMS based magnetometers, however, can be considered as more exotic and less widespread. Therefore, they are not discussed here in more detail.

Magneto-resistive sensors

The amplitude of magneto-resistive effects at room temperature is negligible in non-magnetic metals and semiconductors within a magnetic field range on the order of few (tens of) mT which is relevant for the vast majority of commercial applications. This also applies to geometrical magneto-resistance effects in semiconductors with high room temperature mobility. Strong resistance changes arise for comparatively large magnetic fields, both in case of Corbino disc geometries and, in particular, in extraordinary magneto-resistance (EMR) devices. The EMR effect, discovered in 2000, is universal for any non-magnetic, high-mobility, and homogeneous semiconductor that is shaped into a disk and equipped with a central metal shunt. However, the MR values become extraordinarily large only for magnetic fields of

several hundred mT and beyond [368]. Despite large expectations after its discovery, the EMR effect has not (yet) found its way into applications, e.g. as future read head technology in hard disk drives. The same fate has happened to the colossal magneto-resistance effect (CMR) that arises in perovskite oxides but requires low temperatures. Therefore, substantial magneto-resistance effects, that are exploited in commercial devices, arise only in ferromagnetic materials or in multilayer structures involving ferromagnetic layers:

- Anisotropic magneto-resistance (AMR) based sensors: Despite being the first discovered MR effect, already in 1857, its usage in commercial magnetic field sensors took until the late 1960s. Since then, however, AMR sensors have become widely used and the strongest competitor to Hall effect based devices. Their working principle relies on the dependence of the resistance, that means of the scattering probability, on the angle between the direction of the current and the direction of the magnetization in a thin layer of ferromagnetic material. Since the magnetization direction is influenced by the external magnetic field, a smart device design allows the measurement of the in-plane magnetic field component. Often, the so-called barber pole arrangement is used in order to linearize the magnetic field dependence. It consists of metallic stripes that are placed at an angle of 45° on a thin film of permalloy (Ni-Fe alloy) to dictate the current direction. Typical $\Delta R/R$ values are on the order of few percent. For further details see for instance Refs. 367, 369, 370.

- Giant magneto-resistance (GMR) based sensors: In 1988, a strong magneto-resistive effect ($\Delta R/R$ up to 70 % at room temperature) was observed in multilayer structures consisting of at least two thin ferromagnetic layers separated by a non-magnetic conducting spacer layer [367, 369, 371, 372]. When both ferromagnetic layers have parallel magnetic moments, the scattering at the interfaces and the resistance are low, while for anti-parallel moments the resistance is high. Several different types of GMR systems have been developed so far. Often and for practical reasons, e.g. to obtain a larger absolute effect amplitude, the current in plane configuration is preferred over the current perpendicular to plane configuration [367, 369]. In case of the spin-valve type GMR sensor, an additional antiferromagnetic layer is used to pin the magnetization of one ferromagnetic layer while the other's magnetic momentum remains sensitive to the external magnetic field. The strongest impact of the GMR technology was in magnetic recording. The high sensitivity of GMR based read-heads allowed to reduce the distance between the magnetic domains which boosted the data storage capacity of hard disk drives (compared to AMR read-heads). For further details see for instance Refs. 367, 369, 370.

- Tunnel magneto-resistance (TMR) based sensors: TMR elements are derived from the spin-valve GMR arrangement in which the non-magnetic conducting spacer layer is replaced by an extremely thin, non-magnetic insulating layer (also referred to as magnetic tunnel junction, MTJ) [373]. When a current is applied perpendicular to the plane, electrons tunnel from one ferromagnetic layer into the other. The probability of this tunneling process depends on the mutual orientation of the magnetic moments. It is high (resistance low) in case of parallel, and low (resistance high) in case of antiparallel magnetization between the two ferromagnetic layers. The breakthrough of MTJs happened in 2004 when very thin, crystalline MgO tunnel barriers could be fabricated which resulted in $\Delta R/R$ values of up to 200 % at room temperature (later even up to 600 %) [373–375]. Since TMR elements clearly outperform GMR devices, they soon replaced them in the read-heads of hard disc drives and in MRAMs (magneto-resistive random-access memory). Since the start of their mass-scale production a few years ago, TMR based magnetic field sensors are also becoming more widespread in consumer electronics (electronic compass), industrial or automotive applications.

The above listed magneto-resistive technologies (XMR) are continuously conquering a larger share in the magnetic field sensor market[4]. In particular for applications with high performance requirements (e.g. electronic compasses), XMR devices have often replaced Hall sensors. The latter, however, still have the by far largest share in the overall magnetic field sensing market since cost is the most critical parameter in many use cases and ease of integration is the key. The most recent TMR technology is part of the roadmap of most manufacturers of magnetic field sensors because of its very high sensitivity, ultra-low power consumption, and small temperature drift [376]. Elaborate discussions on the advantages and disadvantages of all XMR and Hall sensor technologies can be found, for instance, in Refs. 367, 369, 370, 376. In the following sections of this chapter, we will compare graphene based Hall elements only to other Hall transducer materials (comparing apples to apples).

Graphene based magneto-resistive sensors

For the sake of completeness, we want to note that not only the Hall effect but also magneto-resistance (MR) effects can, in principle, be utilized to measure magnetic fields at room temperature with graphene based devices (see Refs. [360–362, 377, 378] and Section 4.5). In graphene, magneto-resistance at room temperature can

4 See, for instance, the recent research study "Magnetic Sensors - Market and Technologies 2017" by Yole Développement.

emerge in two ways: Magneto-resistance can be observed either around the CNP due to the coexistence of electrons and holes (originating from graphene's peculiar band structure without band gap) or it arises from geometric effects that can be enhanced in extraordinary magneto-resistance (EMR) device designs (see Refs. [360, 361, 368, 377] for details). In the latter case, the obtainable MR values increase with larger charge carrier mobilities and depend strongly on the applied magnetic field. This makes high-quality, hBN-encapsulated graphene with its record room temperature mobility, even excelling InSb, the potential material of choice for such EMR devices [360, 361]. However, i) the strong magnetic field dependence leads to a non-linear response, ii) very large sensitivity values (defined by dR/dB) are reached only for large magnetic fields beyond one Tesla, and iii) the magnetic field resolution is only on the order of several mT [360, 361]. Therefore, both graphene based as well as in general EMR devices offer much less commercial use than Hall sensors and thus are not discussed further in this work. Finally, the occurrence of the classical magneto-resistance at the CNP is the only material intrinsic magneto-resistance effect at room temperature in graphene. It was already discussed elaborately in Section 4.5 of the preceding chapter: Given the comparatively small magnetic field values that are of interest in commercial applications, the obtainable sensitivity values are small, the magnetic field dependence is quadratic, and several other technical difficulties hamper the exploitation of the magneto-resistance at the CNP for magnetic field sensing. The latter disadvantages motivate our focus on graphene Hall elements in this work.

5.2. Hall sensors and their key performance indicators

Hall effect based magnetic field sensors still have the biggest share in the market and are widely used in automotive, industrial, and consumer electronics applications, as discussed in the preceding section and in Ref. 370. Virtually all commercial Hall sensors are fabricated as integrated devices. They consist of the Hall transducer, the actual sensing part of the device also referred to as Hall element, plate, or generator, and a multitude of other integrated electronic components. Figure 5.1 shows a simplified functional block diagram containing all main components of a typical Hall sensor die[5]. In the following, we will discuss the most important performance indicators of Hall devices and how these other integrated components help

5 In practice, a Hall sensor IC is even composed of many more components that are not included in this simplified picture. Some examples are supply level detection (input voltage range) and stabilization, protection circuits, internal oscillator as clock, programming unit, A/D conversion, or digital signal processing and output interfaces.

to improve the properties of the Hall transducer through sophisticated offset and noise canceling, thermal drift compensation, and linearization. This list of key performance indicators can be found in every data sheet of commercial, integrated Hall sensors where, however, the performance on system level is stated. Therefore, one cannot compare values in these data sheets directly to measurements performed on Hall transducers without any on-chip optimization (as conducted in this work).

5.2.1. High sensitivity at low power consumption

Definitions and general design considerations

Designing a Hall transducer is a trade-off between power consumption and signal amplitude, i.e. the absolute magnetic field sensitivity defined as $S_A = U_H/B$ (in units of V/T). In practice, a fixed supply voltage U is typically given which determines, together with the resistance between the two current leads, the current through the Hall transducer, $I = U/R$, if no additional pre-resistor or sophisticated biasing unit are employed. This is referred to as voltage mode operation and the voltage related sensitivity, $S_V = S_A/U$ in units of V/VT, then constitutes an important performance indicator. From a physical point of view, the Hall effect is a result of the Lorentz force acting on drifting charge carriers. Hence, it is consequential to characterize a Hall transducer also by the current related sensitivity for a constant bias current, $S_I = S_A/I$ in units of V/AT or Ω/T. The latter is equivalent to the Hall coefficient R_H definition in the preceding chapters which there was sometimes also referred to as sensitivity, meaning the current related sensitivity. State of the art Hall sensor ICs are typically equipped with a biasing unit which controls the applied current through the Hall transducer. In this regard, the current related sensitivity is a good performance indicator for the Hall transducer part alone, while the voltage related sensitivity is relevant for the Hall sensor IC as an entity since the supply voltage drives all components on the chip. Therefore, we will focus on the current related sensitivity definition in the subsequent sections studying graphene based Hall transducers without optimization through on-chip components.

Assuming current mode operation and that no other electronic components than the Hall transducer itself have to be considered when calculating the overall power consumption, the following two equations must be optimized:

$$S_A = S_I\, I = G_H(L/W)\,\frac{I}{|q|\, n_s} \qquad P_{el} = U\, I = R\, I^2 = (R_c + \rho\,\frac{L}{W\, d})\, I^2 \tag{5.1}$$

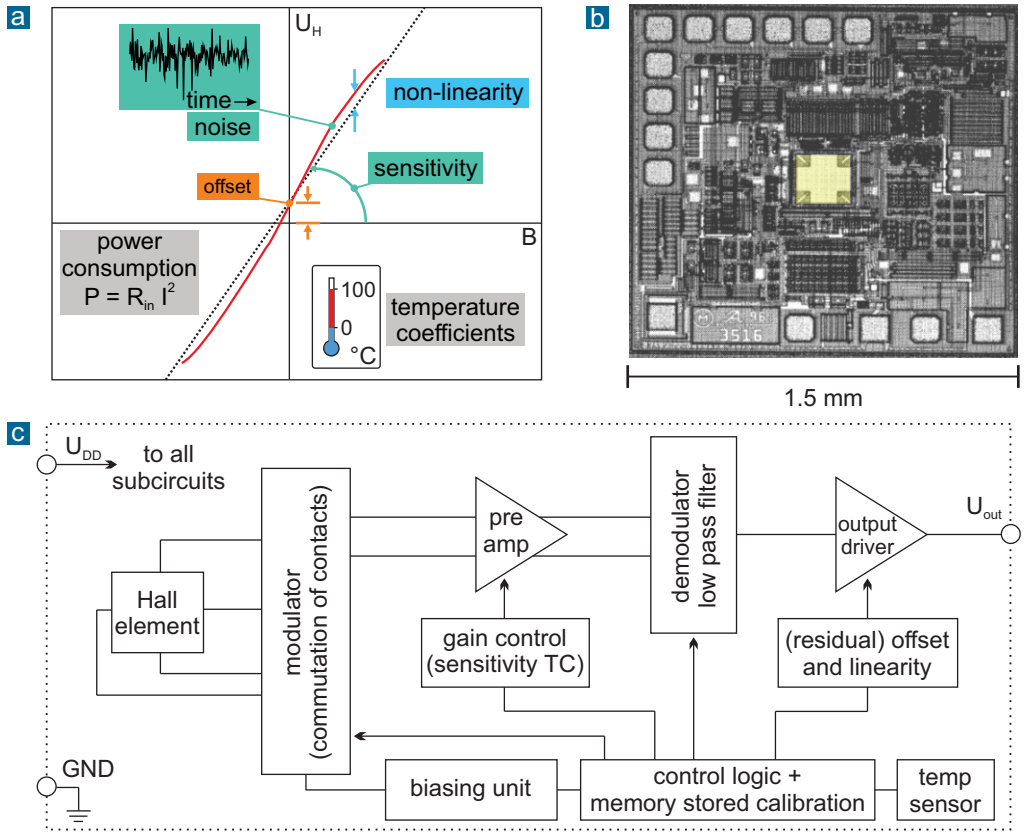

Figure 5.1. Key performance indicators and simplified functional block diagram of a state-of-the-art Hall sensor IC. a: Exemplary $U_H(B)$ output curve of a Hall transducer (red solid line) and illustration of its main characteristics, as discussed in the main text. High sensitivity and low noise level result in a high magnetic field resolution, i.e. small B_{min} value. Input resistance and bias current determine the power consumption. b: Microphotograph of a Si Hall sensor die (taken from Allegro Microsystems website). It demonstrates the high degree of miniaturization and complexity of modern Hall sensor ICs resulting in a restriction of the available area for the Hall plate, the actual sensing part (the highlighted Hall cross in the center measures 160 x 160 μm). c: Simplified functional block diagram of the main components which are integrated on-chip to improve the bare Hall element's properties illustrated in panel a (inspired by several block diagrams, e.g. of the Allegro A1324 Hall sensor). Commercial data sheets typically state the properties after optimization. Then, power consumption has to be considered and minimized at the system level. Modulator, demodulator, and low pass filter are the key elements of the spinning current technique which reduces offset and low frequency noise substantially (see main text).

Here, $n_s = n\,d$ is the sheet carrier density of the Hall element. The length L, width W, and film thickness d (for conventional bulk materials) define the channel geometry

between the current leads. For 2D materials, the thickness variable d is discarded. In the Drude model, the resistivity is given by $\rho = 1/(en\mu)$. Moreover, the contact resistance R_c has to be taken into account for the electrical power dissipation P_{el}. It depends on the effective contact area, given by the width of the metallic contact pad, W_c, and the transfer length between the transducer and the contact pad materials. Since the width is often orders of magnitude larger than the transfer length, e.g. for metallic contacts on 2D materials, the contact resistivity ρ_c is typically stated in units of $\Omega\,\mu$m and the contact resistance can be calculated using $R_c = 2\,\rho_c/W_c$. Finally, $G_H(L/W)$ is the geometrical correction factor of the Hall voltage and is a function of the L/W ratio, as shown in Ref. 150. It approaches unity for $L/W > 3$, i.e. leading to maximum absolute sensitivity, while the optimum between sensitivity and power dissipation is obtained for $L/W \approx 1.3 - 1.4$ (then $G_H(L/W) \approx 0.7$).

From the above equations (Eq. 5.1) follows that the absolute sensitivity is maximized when a large current is applied through a rectangular channel geometry with $L/W > 3$ (such that $G_H(L/W) \to 1$) and when a material with low charge carrier sheet density, i.e. a small charge carrier density and/or a very thin film, is used. However, this results in a larger resistance and the power dissipation scales quadratically with the applied current. Moreover, as we will discuss below and in detail in the following sections, a larger resistance increases the thermal noise floor of the system and the flicker (or 1/f) noise voltage scales linearly with the applied current. The latter eventually leads to a saturation of the magnetic field resolution even when increasing the bias current further. Materials with high mobility at low charge carrier densities constitute the best candidates for Hall transducers since resistance values and power consumption remain, in comparison, moderate[6].

In practice, Hall elements are equipped with four electrical leads, two current contacts and two sense contacts for measuring the Hall voltage. Geometries are chosen to be 90° rotation-symmetric. This allows for a commutation of the four electrical leads which is the pre-requisite for the offset and noise reduction technique discussed below. As a consequence, the input resistance (the resistance between the current leads) and the output resistance (the resistance between the sense leads) are equal, when we presume a homogeneous quality of the transducer material and of the contact resistances[7]. Typical geometries in commercial devices range from Hall cross ($L/W \approx 2 - 3$) to square shapes ($L/W = 1$). In all cases, and in particular in case of Hall cross geometries, the width of the electrical contact pad is (almost) equal to

6 This is also directly reflected in the voltage related sensitivity which scales with the mobility $S_U \propto \mu$ [242].

7 While the power consumption is directly connected to the input resistance, the output resistance is of importance for the layout of the subsequent components on the Hall sensor chip, e.g. in terms of impedance.

the channel width ($W_c \approx W$). The vast majority of our samples in this work, from which we obtained for instance the Hall coefficient statistics shown in Figure 4.2 of the preceding chapter, had a cross shape with $L/W \geq 3$ (such that the geometrical correction factor can be neglected). According to the calculations performed in Ref. 150, $G(L/W) \approx 0.65$ for square shaped devices which one would have to consider when transferring our results to other geometries.

For conventional bulk materials such as Si, the transducer material's thickness has a direct impact on the sensitivity (the thinner the higher) and the input resistance (the thinner the larger). For 2D materials, at first glance, only the L/W ratio appears directly in Eq. 5.1 but no absolute dimensions of the Hall element play a role. This is true for the sensitivity. However, scaling the Hall transducer will lower the contribution of the contact resistance to the overall two-terminal resistance, since $W_c \approx W$ and so $R_c \propto 1/W$. This reduces the power dissipation until the contact resistance becomes negligible compared to the resistance of the channel. Furthermore, as will be discussed in the following section, the absolute dimensions and the bias current primarily determine the noise level in the Hall transducer[8]

Technical measures for improvement

The bias current is mainly limited by power consumption requirements (and the acceptable noise level). As a critical parameter, it has to be chosen carefully and, in general, should be kept as low as possible to achieve the desired performance. In modern Hall sensor devices, the current is applied only during a short period of time, i.e. in form of square wave pulses, which reduces the effective current value on time average. The signal is then enhanced through boxcar averaging. Moreover, the electrical contacts are commutated in order to suppress Hall voltage offset and low frequency noise, as explained below. We refer to the frequency of this commutation process in the following as switching frequency. From flicker noise perspective, it is quasi equivalent to a real AC operation of the device. However, the larger this frequency has to be, in order to effectively reduce the 1/f noise contribution, the more complex and power consuming are the surrounding electrical components.

8 A larger input resistance due to a higher contact resistance contribution increases the thermal noise floor. The 1/f noise scales inversely with the total number of charge carriers in the channel, that means with the volume (or area). Increasing the footprint suppresses 1/f noise and thus improves the signal to noise ratio, if the device is operated in the 1/f noise regime. It extends the possible range for other parameters, such as the bias current, without the 1/f noise becoming dominant. This dependence on the channel area will be demonstrated in the following sections.

Therefore, the Hall sensor IC as a whole has to be taken into account when assessing power consumption and cost of commercial devices.

Summarized, the absolute sensitivity $S_A = U_H/B$ of a Hall transducer is determined by its geometry, the sheet charge carrier density of its material, and the maximum available bias current taking into account power consumption constraints. Moreover, the Hall voltage output for a given external magnetic field can be increased in two ways on Hall sensor chip level. An electronic pre-amplifier can be employed and/or the magnetic flux through the Hall element can be enhanced by means of a magnetic field concentrator [347, 379]. The latter approach is less common in commercial devices since it requires the use of a ferrite substrate and the placement of a ferrite chip on top of the Hall transducer. It offers, however, the advantage that only the Hall voltage signal but not the electrical noise level is increased. In contrast, the use of on-chip pre-amplifiers is standard. The Hall voltage as well as the electrical noise of the Hall transducer at the input stage are amplified equally such that the signal to noise ratio, i.e. the magnetic field resolution, remains unchanged.

5.2.2. High magnetic field resolution and low offset

In real devices, the Hall voltage for a given magnetic field and absolute sensitivity of the Hall transducer, $U_H(B)$, is overlain by spontaneous and random voltage fluctuations that change in time and whose amplitude may depend on the operating frequency, i.e. the noise voltage $U_n(t)$. Therefore, in many applications, not just the absolute sensitivity alone but also the signal to noise ratio, i.e. the magnetic field resolution, is subject to optimization. Moreover, a parasitic ohmic offset voltage, U_{off}, is typically measured in the absence of an external magnetic field between the two Hall voltage contacts of a biased Hall element. If not accounted for, this would be interpreted erroneously as a non-zero magnetic field which is also referred to as offset-equivalent magnetic field, $B_{off} = U_{off}/S_A$. In total, the voltage output of the Hall transducer, can be described as follows:

$$U_{out}(B,t) = U_H(B) + U_n(t) + U_{off}\,. \tag{5.2}$$

The proper Hall voltage, in particular for static magnetic fields, is the harder to discriminate the smaller the external field and the larger the offset and noise voltages are. The offset voltage can be considered as static within the typical time span of a measurement and as long as the environmental conditions and operating parameters

remain stable[9]. It typically exhibits an ohmic characteristic that may originate from a small misalignment of the Hall voltage sensing leads [150, 347]. The latter then pick up a small portion of the longitudinal voltage drop that scales linearly with the bias current. Another source are inhomogeneities in the transducer material itself, such as defects or a non-uniform doping density. Moreover, piezoresistive effects due to mechanical stress in the wafer, induced by the die packaging process, can play a role [150, 347, 380].

In most cases, however, this static offset can be eliminated almost entirely when the Hall plate geometry is designed to be 90° rotation-symmetric. An orthogonal coupling of several Hall transducers fabricated on one and the same chip, that means the use of several interconnected Hall plates with 90° rotated current direction, eliminates the offset contribution when the source of the offset voltage is constant along the chip [379, 380]. This ‚spatial averaging ‘approach, however, consumes much costly footprint on the wafer and increases power dissipation. Therefore, a dynamic offset cancellation technique has been proposed as an alternative, known as the spinning current technique or switched Hall method [379, 381–383]. Only one transducer is used but instead its four electrical leads are commutated with a certain switching frequency which allows for averaging the output signal of the four different biasing configurations. This ‚time averaging ‘offers the advantage that also the residual, transducer specific offset, e.g. due to production spread and wafer inhomogeneities, is reduced[10]. Moreover, this dynamic technique reduces also the low frequency noise, as discussed further below. It thus has become the standard offset (and noise) cancellation technique in modern devices, even though requiring additional components for signal conditioning and treatment in contrast to the static offset cancellation (see also Figure 5.1).

Let us now focus on the noise voltage, $U_n(t)$, which we assume to be an analog, small-signal perturbation. While the time average of the noise voltage is typically zero, it will be non-zero for the noise power for which the mean-square noise voltage is typically chosen as descriptor (since it is proportional to the power dissipation). The latter is defined as follows [150, 384]:

9 Since typically having an ohmic contribution, the offset voltage scales with the applied current and drifts with temperature. Moreover, defects in the Hall transducer material, e.g. near the contact areas, may migrate along the device channel with time and thereby slowly alter the offset value. These drifts in the offset voltage should be as small as possible in order to achieve a stable device performance. However, they are less critical when offset cancellation techniques such as the spinning current method are applied.

10 Still, a small residual offset can remain which is normally negligible for most commercial applications. This sensor individual offset can be only mitigated through calibration which is expensive and hence only performed when ultra high precision is needed.

$$\overline{U_n^2} = \lim_{T\to\infty} \frac{1}{T} \int_0^T [U_n(t)]^2 \, dt \; . \tag{5.3}$$

The root mean-square (RMS) noise voltage is then given by $U_{n,rms} = (\overline{U_n^2})^{1/2}$. As discussed in detail in the next section, some noise types can exhibit a frequency dependence. Therefore, it is common to determine and state the noise power spectral density, $N_{PSD}(f)$ in units of V^2/Hz, which is represented by the mean-square noise voltage per bandwidth Δf (since the latter is proportional to the power dissipation per frequency range of a resistor). Often, also the (RMS) noise voltage spectral density, $N_V(f)$ in units of $V/\sqrt{Hz}$, defined as the RMS noise voltage per bandwidth, is given. The mathematical definitions read as follows [150, 384]:

$$N_{PSD} = \lim_{\Delta f\to 0} \frac{\overline{U_n^2}}{\Delta f} \qquad N_V = \lim_{\Delta f\to 0} \frac{U_{n,rms}}{\Delta f} \; . \tag{5.4}$$

Both quantities can be easily converted into each other via $N_V(f) = \sqrt{N_{PSD}(f)}$. Modern spectrum analyzers provide at least one of these two quantities as output value over frequency. Hence, they perform the limit value calculation within the set resolution bandwidth (RBW) automatically. The resolution bandwidth should be chosen in accordance to the number of data points per frequency range such that the resulting values are not too much smeared out over the respective frequency interval[11]. We note that the noise voltage spectral density and sometimes even the noise power spectral density are quite often referred to as "noise voltage" in the literature (a rather sloppy notation which may lead to confusion, in particular when no units are provided).

When independent and uncorrelated, different noise sources add normally to the total noise power spectral density (in mean-square units) but add in root sum of squares manner to the total noise voltage spectral density (in root mean-square units):

$$N_{PSD,tot} = N_{PSD,therm} + N_{PSD,1/f} + N_{PSD,amp} + ... \tag{5.5}$$

11 For instance, we have used the following settings: scanning of the frequency range 0 to 1 MHz in 1001 steps (that means 1 kHz intervals) with RBW set to 100 Hz. Range 0 to 50 kHz in 2001 steps (that means 25 Hz intervals) with RBW set to 10-20 Hz.

$$N_{V,tot} = \sqrt{N_{V,therm}^2 + N_{V,1/f}^2 + N_{V,amp}^2 + ...} \tag{5.6}$$

In the above equations, we have already considered only the most common contributors to electrical noise in Hall transducers. Intrinsic sources, that means noise generated by the transducer itself, are mainly thermal noise and flicker (1/f) noise, as discussed in more detail in the following section. Moreover, the pre-amplifier constitutes an additional noise source, determined mostly by its input noise level (see Section 5.4 on the experimental setup for details). After this amplification step, further environmental or instrumental noise sources typically do not play a role anymore since the amplified Hall voltage clearly exceeds their amplitudes.

Analogous to the above definition of the offset-equivalent magnetic field, one can define the noise-equivalent magnetic field. Other common synonyms in data sheets or in the literature are magnetic field resolution, detectivity, or minimum detectable magnetic field, since this parameter reflects the magnetic field value for which the Hall voltage signal equals the noise voltage. It depends on both the absolute sensitivity and the noise level, and thus constitutes the most important performance indicator of a Hall device. Again, it is common to state its spectral density in units of $T/\sqrt{Hz}$, taking into account a possible dependence on the frequency. It reads as follows [348]:

$$B_{min} = \frac{N_{V,tot}}{S_A} = \frac{\sqrt{N_{V,therm}^2 + N_{V,1/f}^2 + N_{V,amp}^2 + ...}}{S_I\, I} \tag{5.7}$$

In order to obtain a high magnetic field resolution, that means a small B_{min} value, the absolute sensitivity has to be high and the noise level low. The conditions for a high sensitivity in current mode operation have already been discussed above: small carrier density, large bias current and proper channel length to width ratio (see Eq. 5.1). However, the conditions for a low noise level are not universal (see also Section 5.3). The latter will depend on the intrinsic properties of the transducer material, contact resistances, specific device fabrication steps that may potentially degrade quality, given environmental and operating conditions as well as constraints on the maximum footprint of the transducer. In general, an optimization of the processing steps, which result in an improved quality of the transducer material and the electrical contacts, will also reduce the noise level until eventually material intrinsic limits are reached. The latter are then determined by the dominating noise

mechanism for the given operating parameters, e.g. bias current or frequency, as well as environmental conditions, i.e. mainly temperature.

The noise characteristics of a Hall transducer typically comprises of an interplay between thermal and flicker (1/f) noise. Flicker noise will prevail over the frequency-independent thermal noise floor at small enough frequencies, i.e. below the corner frequency f_c. The latter typically ranges between 100 Hz and 100 kHz (see Refs. [379, 380]). The strength of the 1/f noise and thus the corner frequency increase with the bias current but decrease with the total number of charge carriers, given by channel area times sheet carrier density. From the latter follows that a larger device area (or volume for bulk materials) is beneficial. However, the demand for miniaturization in modern devices limits the available footprint for the transducer, the actual sensing part of a Hall sensor IC, to a typical range between 100x100 μm^2 and 1000x1000 μm^2 (see also Figure 5.1). Summarized, the electrical noise of the Hall transducer is a complex function of the two-terminal resistance and temperature, determining the thermal noise floor, as well as carrier density, channel area, bias current and operating frequency, determining the 1/f noise contribution (see Section 5.3 for details). Moreover, taking into account the functional dependencies of the absolute sensitivity on the bias current and carrier density adds further functional complexity. As will be discussed in more detail in the following section, the current dependencies in the nominator and denominator of Eq. 5.7 cancel out if 1/f noise prevails. Therefore, any further increase of the bias current will not improve the magnetic field resolution anymore although the absolute sensitivity increases. The B_{min} value saturates with respect to the current in the 1/f noise regime.

To conclude, the main task is to find the ideal working point of the Hall transducer under certain application specific constraints such as available area, operating frequency, and power consumption. This means balancing sensitivity and noise that have contrary requirements for certain parameters such as the bias current. Although the use of a pre-amplifier makes the output signal more robust against other environmental or instrumental noise, it will not improve the magnetic field resolution since both Hall signal and noise are amplified to the same extent. Moreover, the amplifier's contribution to the total noise level has to be negligible which demands the employment of ultra low noise pre-amplifiers (see Eqs. 5.6 and 5.7). The spinning current technique, which was already introduced as a dynamic method to minimize the offset voltage, has been found to reduce also substantially the low frequency noise while causing only a negligible increase in the white noise floor [379, 380, 383, 385]. The simplified explanation is that the noise voltage $U_n(t)$ in Eq. 5.2 can be interpreted as a fluctuating offset voltage. In other words, the low frequency noise voltage can be considered as generated by a fluctuating offset resistance times the bias current [383]. Accordingly, the time averaging while commutating the contacts

will cancel out not only the static offset voltage but also the noise voltage components in the frequency spectrum notably below the switching frequency since these components then appear static, too. When the switching frequency is substantially higher than the corner frequency, the 1/f noise contribution is removed almost completely. Hence, typical switching frequencies are in the range between 10 kHz and up to several 100 kHz [383, 385]. In general, the higher the frequency, the more complex and power consuming are the integrated components for signal modulation and processing. This has to be taken into account when assessing the performance on the sensor level (see Section 5.8). However, since the spinning current technique substantially improves the performance of the bare Hall transducer, operating frequencies of even up to 1 MHz are expected to be feasible and could be implemented from a technical point of view when necessary.

5.2.3. Other important characteristics

Small linearity error

Ideally, one expects the output Hall voltage to show a perfectly linear dependence on the external magnetic field. However, several effects can lead to a deviation of the actual output characteristics from such ideal linear response. Non-linearity can originate from geometrical effects, a magnetic field dependence of the Hall coefficient, $R_H(B)$, a large electric potential difference between the contacts due to high bias voltage, or a doping inhomogeneity along the Hall plate [150]. Moreover, saturation of the output voltage will eventually occur at high magnetic fields since Hall transducers are passive devices and the Hall voltage cannot exceed the bias voltage [347]. In general, high mobility transducer materials and large magnetic fields favor non-linearity. Therefore, for each Hall sensor a maximum magnetic field range is provided in its data sheet. The stated maximum linearity error then refers to this range. Both the material non-linearity, i.e. the strength of the magnetic field dependence of the Hall coefficient, as well as the geometrical non-linearity, i.e. the strength of the magnetic field dependence of the geometrical correction factor, are proportional to $\mu_H^2 B^2$ but exhibit opposite signs [150, 380]. This allows designing devices where these two effects almost cancel each other. Moreover, geometrical non-linearity can be compensated by loading the output with a proper resistor [150].

Our devices in this work were not optimized in any particular manner in order to reduce the linearity error. The used Hall cross geometry and the L/W ratios are

common in commercial devices and, thus, our results comparable to reported linearity errors elsewhere. However, the peculiarity of graphene compared to other Hall transducer materials is the absence of a band gap and the coexistence of electrons and holes near the CNP, as discussed in detail in the preceding chapters. Therefore, the non-linearity due to the magnetic field dependence of the Hall coefficient will increase at low charge carrier densities when both carrier types coexist (see Eq. 2.61 in Section 2.3.3). This linearity error also has a detrimental impact on the accuracy when determining the Hall coefficient by means of a linear fit to the $U_H(B)$ data points. Therefore, we already restricted the maximum magnetic field range in the previous chapters such that $\mu^2 B^2 < 1$ holds. Depending on the charge carrier mobility of our devices, this corresponds to a magnetic field range on the order of $50 - 100\,\mathrm{mT}$ which is considered for the analysis (see also Section 5.4). Within this range, non-linearity remains fairly small even at the CNP. This studied magnetic field range is comparable to the maximum magnetic field ratings of most commercial Hall sensors (in case of geomagnetic sensors, maximum ratings are on the order of few mT only). Hence, our measurement conditions are realistic.

In contrast to the other key performance indicators, there is no consistent definition of the linearity error throughout all literature and commercial data sheets. The straightforward approach is to define the non-linearity as the percentage deviation of the measured Hall voltage for a certain bias current and magnetic field, $U_H(B, I)$, from the linear fit to the measured values, $U_{H,fit}$ [150, 380]. The non-linearity then reads as follows:

$$NL = \frac{U_H(B, I) - U_{H,fit}}{U_{H,fit}} \cdot 100\% \tag{5.8}$$

This definition, however, does not provide one single value to assess the linearity over the entire specified magnetic field range of the sensor. For instance, the deviations in percent according to Eq. 5.8 can be expected to be higher at small Hall voltages (for small magnetic fields) since then also other effects such as the zero field offset and noise can play a role. Instead, the full-scale non-linearity is provided in most commercial data sheets which sets the deviations in relation to the full-scale Hall voltage output at the specified maximum magnetic field (see, for instance, the data sheet of the BMM150 sensor from Bosch Sensortec). There are also different ways to define the linear fit to the measured values. In case of the end point method, the non-linearity is defined as the maximum deviation from a straight line between the two end points, relative to the full-scale output. More common, however, is the best fit straight line method (BFSL) which typically uses a least squares algorithm for

the linear fit to all data points and then normalizes these deviations to the full-scale output value. In our case, we use the polyfit function in Matlab to obtain the best fit straight line and then divide the deviation by the full-scale Hall voltage output at maximum magnetic field. Since this is computed for every data point, both the mean as well as the maximum full-scale linearity error can be provided according to the following equation:

$$\mathrm{max/mean}(\mathrm{NL_{FS,BFSL}}) = \mathrm{max/mean}\left(\frac{\mathrm{U_H(B,I)} - \mathrm{U_{H,BFSLfit}}}{\mathrm{U_{H,FS}}}\right) \cdot 100\% \quad (5.9)$$

The maximum full-scale non-linearity, i.e. $\mathrm{max}(\mathrm{NL_{FS,BFSL}})$, is typically about 1% in commercial devices (e.g. for the BMM150 magnetic sensor from Bosch Sensortec). Aside from the above mentioned design optimization and maximum magnetic field restrictions, linearity errors are difficult to remove. However, calibration can be performed and stored in on-chip memory. The output curve is then linearized through software based corrections.

Small temperature drift

Like every sensor, Hall devices suffer from an undesirable cross-sensitivity to other environmental parameters, for instance pressure/mechanical stress or temperature (see Refs. 150, 347, 380). In particular, temperature drifts are prevalent and critical, and thus specified in every data sheet. Both the mobility as well as the charge carrier density depend on temperature. Moreover, thermally induced mechanical stress may cause piezoresistive effects. Therefore, the sensitivity (and with it the magnetic field resolution), the offset voltage and the resistance of the Hall transducer drift over temperature. Considering the Hall sensor on system level, also other integrated components will exhibit a temperature dependence. This may lead to a rather complex thermal behavior of the Hall sensor. Therefore, a maximum operating temperature range is specified in each data sheet for which the given temperature coefficients are valid, e.g. -40 to 85°C (125°C) for Hall sensors utilized in consumer (automotive) applications.

Hall transducer materials with small band gap, such as InSb or graphene as the most extreme case with zero band gap, and low charge carrier density will be affected mostly from thermal activation of charge carriers. On the contrary, high band gap materials, such as Si or GaAs, and highly extrinsic materials exhibit a smaller

temperature drift of the Hall coefficient. Moreover, the resistance change is governed by the temperature dependence of both density and mobility (which depends on the prevailing scattering mechanism in the specific material). This resistance drift is critical in several ways. First, assuming a constant bias voltage and no further compensation, the bias current changes and with it the absolute sensitivity, i.e. the Hall voltage output is different even though the external magnetic field remains the same. Second, in case the offset voltage originates from a small misalignment of the Hall voltage sensing leads, its value will also change with the resistance and bias current[12]. Finally, a high stability of the input and output resistance is also desirable from a circuit design point of view.

Technical measures allow for a compensation of temperature drifts to a large extent [150]. For example, a temperature sensor can be integrated on the chip and serve as input for adjusting the operating parameters in order to obtain a temperature stabilized sensitivity and offset voltage (see Figure 5.1). Typically, modern Hall sensors are equipped with a sophisticated biasing unit which controls precisely the applied current to the transducer. The respective calibration curves are stored in on-chip memory. Moreover, the gain of the pre-amplifier can be adjusted slightly via a central control logic such that the output voltage of the Hall sensor IC remains constant even though the Hall voltage output of the transducer is subject to temperature drifts (see Figure 5.1).

In Section 4.3 of the preceding chapter, we already studied the temperature dependence of the Hall coefficient and the resistivity in graphene devices over a wide temperature range between room and liquid helium temperature using a cryostat system. Our custom-built measurement setup, we describe in detail in Section 5.4, allows studying all aforementioned key performance indicators of Hall elements and provides the capability to set the temperature between room temperature and up to 80°C in a controlled fashion. In the following, we provide for several temperature steps experimental data of noise, sensitivity, resistance, linearity error and offset as a function of density. From the latter, the temperature drifts can be evaluated (see Section 5.7).

12 In addition, the aforementioned change in the piezoresistivity due to thermally induced mechanical stress may affect the offset voltage. In general, devices with larger initial offset also tend to show higher offset drifts [347].

Cost competitiveness

Even though the cost of a Hall sensor is technically no performance indicator, we want to emphasize here again the importance of cost competitiveness in commercial applications. This imposes constraints on the choice of the Hall transducer material, operating parameters such as frequency, as well as complexity of the integrated components (see also Section 5.8). For scalable production processes, the fabrication and handling of the transducer material has to be both cheap and compatible to Si CMOS technology (ideally monolithic integration). MBE growth, for example, is more expensive than other growth processes such as CVD. Moreover, when the intrinsic performance of a Hall transducer material is high, the requirements on the other integrated components are less demanding which results in lower complexity. For instance, when the intrinsic low frequency noise level is low, a further reduction by means of the spinning current technique is either not necessary or a smaller switching frequency can be used.

5.3. Introduction to noise in graphene devices

5.3.1. Noise in electronic devices

All electronic components have inherent noise which means a random, unwanted fluctuation in the electrical signal. An example is the noise voltage component, $U_n(t)$, of a Hall transducer's output signal we described in Eq. 5.2 in Section 5.2.2. The noise frequency spectrum varies substantially among different electronic devices as it may originate from several different effects. Even among different samples of the same type of electronic device and same material system, e.g. graphene Hall transducers, the amplitude of certain noise sources such as the flicker noise depends on the specific device fabrication and processing steps. In summary, the noise spectrum depends on the electronic device type, the material system as well as the actual fabrication and processing steps. Moreover, we consider in the following only intrinsic noise sources, i.e. noise generated by the electronic device itself, but not extrinsic noise sources such as environmental noise (e.g. electromagnetic interference). For integrated devices or several electronic devices connected in series, the noise spectrum of the measured signal is a combination of each component's noise. For instance, in our measurement configuration described in the following section, the actually measured noise voltage composes of the Hall element's intrinsic noise, the amplifier's noise and the noise of the spectrum analyzer. If different noise sources

are independent and uncorrelated, the total noise voltage spectral density (in units of $\mathrm{V}/\sqrt{\mathrm{Hz}}$) equals the root sum of the squares of each noise source (see Eq. 5.6 in Section 5.2.2).

Thermal noise

In this work, we focus on the common noise sources in analog devices that have no gap or potential barrier[13], such as Hall elements. Such devices then constitute a simple resistor with respect to noise. In this case, the white noise is governed solely by thermal noise which is also known as Johnson-Nyquist noise[14]. Thermal noise results from the random motion of charge carriers at finite temperature and is ubiquitous in all resistive elements regardless of applying a voltage [384]. The spectral density of thermal noise exhibits no frequency dependence (white noise). It cannot be reduced by any measure or technique except for lowering the temperature [150, 386]. Thermal noise ultimately defines the noise floor, i.e. the lowest possible noise level that can be reached after eliminating all other noise sources or in the high frequency regime when 1/f noise becomes smaller than the thermal noise floor. The noise voltage spectral density of the thermal noise is calculated as follows (in units of $\mathrm{V}/\sqrt{\mathrm{Hz}}$)

$$\mathrm{N_{V,therm}} = \sqrt{\mathrm{N_{PSD,therm}}} = \sqrt{4\,\mathrm{k_B}\,\mathrm{T}\,\mathrm{R}}\,. \tag{5.10}$$

The resistance between the two electrical contacts that provide the output signal, e.g. the Hall contacts in case of Hall transducers, is composed of the channel resistance and the contact resistance. For 2D materials and using the Drude model for the channel resistance, the resistance is given by

$$\mathrm{R} = \mathrm{R_{ch}} + \mathrm{R_c} = \frac{1}{\mathrm{e}\,\mathrm{n_t}\,\mu}\frac{\mathrm{L}}{\mathrm{W}} + 2\,\frac{\rho_\mathrm{c}}{\mathrm{W}}\,. \tag{5.11}$$

13 A detailed discussion of other noise sources in digital and analog devices is provided, for instance, in Ref. 384.

14 Shot noise contributes only substantially to the white noise for devices with a gap or potential barrier, i.e. in case of semiconductor pn junctions in diodes or MOSFETs, since it originates from the discrete nature of the charge carriers which causes a random fluctuation of the current [384].

In the following we consider 90° rotation-symmetric cross or square shaped Hall elements with channel length L and width W. In this case, the input resistance (between the source and drain contacts) and the output resistance (between the Hall contacts) are equal if the contact resistances of all contacts are almost identical. The latter constitutes a fair assumption given that all contacts of the device are fabricated in the same process. Moreover, we measured the resistance between all contact pairs of each Hall element studied in this work and sorted out devices that did show a notable deviation among the contact pairs. ρ_c is the contact resistivity in units of $\Omega\,\mu$m which can exhibit a density dependence. Its importance will be discussed in more detail in Section 5.7.1. ρ_c is also one of the four key parameters we use to model graphene Hall devices (see Section 5.7.2). In case of graphene where electrons and holes coexist near the CNP, n_t corresponds to the total charge carrier density which can be determined using the channel approximation (see Eq. 2.30 in Section 2.2.4). This introduces the residual total charge carrier density at the CNP, $n_{t,0}$, as another parameter for the simulation:

$$n_t = \sqrt{n_{t,0}^2 + n_g^2}\,. \tag{5.12}$$

1/f noise

Besides the frequency-independent white noise, flicker or 1/f noise (also known as pink noise[15]) is an ubiquitous effect in a large variety of physical systems. It is present not only in electronic devices but also in music, speech and signals in human cognition [389]. Flicker noise usually prevails in the low-frequency regime exceeding other noise sources such as thermal noise by orders of magnitude. Since its power spectral density depends inversely on frequency, there is eventually a frequency value, the so-called corner frequency f_c, for which flicker noise is equal to the thermal noise. Above this corner frequency, thermal noise prevails and defines the floor of the spectral density (as shown in the double-logarithmic plots of N_V versus frequency in Fig.

15 Depending on its definition, pink noise may also include random telegraph noise (RTN) and generation-recombination noise (G-R noise) (see Refs. 384, 386–388 for details). Both are typically present in semiconductor devices with a band gap. G-R noise originates from the random trapping and detrapping of charge carriers. The characteristic relaxation time of this process then relates to a certain frequency in the noise spectrum below which the spectral density is constant and above which it decreases similar to 1/f noise. G-R noise is often observed on top of the 1/f noise spectrum as a bump at a certain frequency. RTN, also known as burst or popcorn noise, results from random but discrete switching events between different, typically two, states of the system.

5.6 in Section 5.5). Flicker noise was studied for several decades, in particular in electronic devices, but its origin is still under debate and related to different mechanisms that depend on the material system and even the actual device fabrication processes [387–391].

Various models, some of high complexity, have been proposed. They are discussed, for instance, in Refs. 387, 388, 392. The two most prominent and common models, the McWhorter and the Hooge model, relate 1/f noise to charge carrier number and mobility fluctuations, respectively [388, 389]. For homogeneous and clean metals or semiconductors such as GaAs, the combination of longitudinal noise measurements, i.e. the measurement of conductance fluctuations, and Hall voltage noise measurements supports the mobility fluctuation model and, accordingly, the Hooge model to be applicable [389, 392, 393]. For other semiconductors such as Si CMOS with potential traps in the oxide, 1/f noise is captured better by the McWhorther model which is essentially based on the summation of many generation-recombination events [389]. As discussed further below in more detail, the Hooge model describes well the 1/f noise in graphene devices. The empirical model provides the following relation for the noise power spectral densities in units of V^2/Hz [387, 388, 390, 392]

$$\frac{N_{PSD,G}}{G^2} = \frac{N_{PSD,R}}{R^2} = \frac{N_{PSD,I}}{I^2} = \frac{N_{PSD,V}}{U^2} = \frac{\alpha_H}{N f^{\beta}}\,, \tag{5.13}$$

where G is the conductance, R is the resistance, V is the voltage and I is the current. N is the total number of charge carriers in the volume involved in the generation of 1/f noise, f is the frequency and α_H is the dimensionless Hooge parameter which gives a normalized estimate of the 1/f noise level in the system. At first, α_H was considered as a universal constant of $2 \cdot 10^{-3}$ but then was found to vary between 10^{-7} and 10^{-2}, depending on the quality of the material and processes involved, such as electrical contacts to the material [388]. In Eq. 5.13, we also considered a frequency exponent for which one typically finds $\beta \approx 1 \pm 0.2$ [392]. It characterizes the 1/f noise further but usually it is very close to unity. Because of that, we will assume $\beta = 1$ in the following in order to reduce complexity in the data analysis and simulation of graphene Hall devices. Since we will determine the 1/f noise level by measuring voltage fluctuations, the relevant parameter is the noise voltage spectral density, $N_{V,1/f} = \sqrt{N_{PSD,V}}$. From the above empirical Hooge relation, it follows that

$$\begin{aligned} N_{V,1/f} &= \sqrt{\frac{\alpha_H}{n_t\, A\, f}} \cdot U \\ &= \sqrt{\frac{\alpha_H}{n_t\, A\, f}} \cdot R \cdot I \\ &= \sqrt{\frac{\alpha_H}{n_t\, A\, f}} \cdot \left(\frac{1}{e\, n_t\, \mu} \frac{L}{W} + 2\, \frac{\rho_c}{W} \right) \cdot I \,. \end{aligned} \tag{5.14}$$

In the above equation we already considered the case for 2D materials where the total number of charge carriers is given by area, $A = L \cdot W$, times total charge carrier density, n_t. Eq. 5.11 is used for the total resistance which includes channel and contact resistances. Using the channel approximation for n_t (Eq. 5.12) completes our empirical model for 1/f noise in graphene. It will be used throughout the following sections, and in particular in Section 5.7.2). In case that the contact resistance is negligible compared to the channel resistance, Eq. 5.14 can be simplified and re-arranged as follows

$$N_{V,1/f}(R_{ch} \gg R_c) = \sqrt{\frac{\alpha_H}{n_t^3\, f}} \cdot \frac{1}{e\, \mu} \cdot \sqrt{\frac{L}{W}} \cdot \frac{1}{W} \cdot I \,. \tag{5.15}$$

Accordingly, we expect a $n_t^{-3/2}$ density dependence of the 1/f noise in graphene devices with negligible contact resistance contribution to the total resistance, in particular in the single carrier regime away from the CNP where the mobility itself exhibits no additional density dependence (in our simulation in Section 5.7.2, we will assume the mobility to be constant).

Finally, the thermal and noise voltage spectral densities, given by Eqs. 5.10 and 5.14, respectively, add in root sum of squares manner to the measured, total noise voltage according to Eq. 5.6. Furthermore, the so-called corner frequency, f_c, can be mathematically defined as the frequency for which $N_{V,1/f}(f_c) = N_{V,therm}$ for given density and current. In the same way a critical current, I_c, can be defined for given density and frequency.

5.3.2. Noise and magnetic field resolution of Hall transducers

The formulas for thermal and 1/f noise introduced in the previous section describe, strictly speaking, only longitudinal resistance fluctuations (which originate from mobility fluctuations in case of the Hooge model). Experimentally, this corresponds to measuring longitudinal voltage or current fluctuations, i.e. in longitudinal configuration between the source and drain contacts. However, for Hall elements, the voltage fluctuations between the Hall contacts, i.e. in transverse configuration, are of relevance and measured in this work. Since we consider 90° rotation-symmetric cross or square shaped Hall elements, the input and output resistances of the Hall transducer are equal (if all contact resistances are approximately equal). Therefore, the thermal noise floor is equal in both longitudinal and transverse configuration.

As already shown in experiments several decades ago, 1/f noise across the Hall electrodes is also governed by conductivity fluctuations originating from mobility fluctuations [383, 393, 394]. Accordingly, the empirical Hooge model is applicable. Moreover, the longitudinal resistance and the Hall voltage fluctuations are correlated in semiconductors [395]. The ratio of the Hall to the longitudinal noise power spectral densities, normalized by the Hall and bias voltage, respectively, is between 0.6 and 1 [396]. For graphene devices on SiO_2, the ratio has been found to be above one, i.e. the Hall voltage noise can exceed the longitudinal noise [395].

In conclusion, Equations 5.14 and 5.15 also apply for 1/f noise measured between the Hall contacts. If not specified otherwise in this work, $N_{V,1/f}$ and $N_{V,therm}$ always refer to the noise voltage spectral density in transverse configuration. It is measured in the absence of a magnetic field for both zero and several different bias current values, as described in detail in the following section. Since the devices are 90° rotation-symmetric and the ratio between Hall and longitudinal noise is approximately unity, the voltage or current in Eqs. 5.14 and 5.15 corresponds to the bias voltage or bias current between the source and drain contacts. This means that the noise voltage spectral density between the Hall contacts can be normalized by the bias voltage, or by the bias current times the two-terminal source-drain resistance in case of constant bias current operation as in our experiments. All other studies on graphene Hall elements also find the noise voltage spectral density between the Hall contacts to scale linearly with the bias voltage or bias current, equivalently, in the 1/f regime and normalize the spectral density accordingly [198, 348, 349, 397–400].

As discussed in Section 5.2 and Subsection 5.2.2, the most important performance indicator of Hall elements is the magnetic field resolution (spectral density), B_{min}

in units of $T/\sqrt{Hz}$. Considering only the most relevant intrinsic noise sources, i.e. thermal and 1/f noise, it is given by (Eq. 5.7 repeated for convenience)

$$B_{min} = \frac{N_{V,tot}}{S_A} = \frac{\sqrt{N_{V,therm}^2 + N_{V,1/f}^2}}{S_I\, I}\,. \tag{5.16}$$

Our model for the magnetic field resolution of Hall elements with two charge carrier types, such as graphene, further includes:

- Eqs. 5.10 and 5.14 for the thermal and 1/f noise voltage spectral density, respectively.
- Eq. 5.11 for the two terminal resistance.
- The two carrier Drude model for the current-related sensitivity, S_I, given by Eq. 4.3 in Section 4.1.2 and Eq. 5.1. For $L/W \approx 3$ (Hall cross geometry), the geometrical correction factor G_H is ≈ 1, while for $L/W \approx 1$ (square) it is about 0.7 [150].
- The channel approximation of the total charge carrier density, Eq. 5.12, which introduces the residual charge carrier density at the CNP, $n_{t,0}$, as ad-hoc parameter.

Let us in the following consider only the single charge carrier case and $G_H \approx 1$. This means that we replace the total charge carrier density by the single carrier density in the above equations and S_I is simply given by $1/(e\,n)$. In the thermal noise floor limit, the magnetic field resolution is calculated as follows:

$$\begin{aligned} B_{min,therm} &= \frac{N_{V,therm}}{S_I\, I} \\ &= \sqrt{4\,k_B\,T\,R}\cdot\frac{e\,n}{I} \\ &= \sqrt{4\,k_B\,T\left(\frac{1}{e\,n\,\mu}\frac{L}{W} + 2\,\frac{\rho_c}{W}\right)}\cdot\frac{e\,n}{I}\,. \end{aligned} \tag{5.17}$$

When the contact resistance is negligible, this simplifies to

$$B_{\mathrm{min,therm}}(R_{\mathrm{ch}} \gg R_{\mathrm{c}}) = \sqrt{4\,k_{\mathrm{B}}\,T} \cdot \sqrt{\frac{L}{W}} \cdot \sqrt{\frac{e\,n}{\mu}} \cdot \frac{1}{I}\,. \tag{5.18}$$

From this follows that the magnetic field resolution in the thermal noise regime is improved, i.e. B_{min} is smaller, at lower densities and when applying larger currents. The thermal noise limit is reached at high enough frequencies above the corner frequency. The corner frequency, however, also shifts to larger frequencies when the density is lowered or the bias current is increased because the 1/f noise amplitude grows (see Eq. 5.15). Alternatively, 1/f noise could be suppressed by the spinning current method introduced in Section 5.2. Since we consider in this work the performance of bare Hall transducers without further optimization on system-level, we refrain from a detailed discussion.

In practice, there will always apply application specific limits in terms of maximum operating frequency or current. The choice of the charge carrier density working point must also be aligned with other requirements, e.g. in terms of the sensitivity, linearity or other key performance indicators. For given operating frequency and density working point, the current can be increased in order to improve the magnetic field resolution until the 1/f noise becomes equal to the thermal noise floor. Beyond this critical current value, I_{c}, only minor improvements of B_{min} are obtained when the current is increased further. Eventually, when 1/f noise prevails, the magnetic field resolution saturates at the lowest possible value that does no longer depend on the applied current. $B_{\mathrm{min,1/f}}$, or $B_{\mathrm{min,sat}}$ used equivalently, can be determined as follows:

$$\begin{aligned} B_{\mathrm{min,1/f}} = B_{\mathrm{min,sat}} &= \frac{N_{\mathrm{V,1/f}}}{S_{\mathrm{I}}\,I} \\ &= \sqrt{\frac{\alpha_{\mathrm{H}}}{n\,A\,f}} \cdot R \cdot I \cdot \frac{e\,n}{I} \\ &= \sqrt{\frac{\alpha_{\mathrm{H}}\,e^2\,n}{L\,W\,f}} \cdot \left(\frac{1}{e\,n\,\mu}\frac{L}{W} + 2\,\frac{\rho_{\mathrm{c}}}{W}\right)\,. \end{aligned} \tag{5.19}$$

Again, when the contact resistance is negligible, this simplifies to

$$B_{\mathrm{min,1/f}}(R_{\mathrm{ch}} \gg R_{\mathrm{c}}) = B_{\mathrm{min,sat}} = \sqrt{\frac{\alpha_{\mathrm{H}}}{f}} \cdot \frac{1}{\mu} \cdot \frac{1}{\sqrt{n}} \cdot \sqrt{\frac{L}{W}} \cdot \frac{1}{W}\,. \tag{5.20}$$

Under the assumption that the mobility and α_H do not or only weakly depend on the density, $B_{min,1/f}$ is lower, i.e. the magnetic field resolution is better, for higher charge carrier density. However, a larger current must then be applied in order to remain unambiguously in the 1/f noise regime. Moreover, the assumption that the contact resistance is negligible not necessarily holds over a wide density range. Hence, the interplay between thermal and 1/f noise is more complex in real devices as will be discussed in the following sections of this chapter.

In order to compare the reported magnetic field resolutions of different Hall elements, it is common practice to normalize the magnetic field resolution by multiplication with the channel width, $B_{min} \cdot W$, and extrapolate the reported values to one reference frequency, e.g. 1 kHz, if they were measured at different frequencies. In doing so, one assumes that the magnetic field resolution was measured in the 1/f noise regime and Equation 5.20 applies. We discuss this in more detail in Section 5.7.3 and propose a definition of $B_{min,norm}$ which also accounts for different geometries through an additional multiplication with $\sqrt{W/L}$.

5.3.3. 1/f noise in graphene

As final part of this introduction, we will briefly review the most important findings of various studies of 1/f noise in graphene devices. Other reviews are provided, for instance, in Refs. 389 and 392. The following summary will focus on noise measurements in longitudinal configuration. This means that longitudinal voltage or current fluctuations were measured revealing the underlying resistance (conductance) fluctuations. Hall voltage fluctuations, i.e. noise in transverse configuration, will be studied in more detail in the following sections which includes also the discussion of other experimental findings in the literature where possible and suitable (for instance, Refs. 198, 348, 349, 397–400). In particular, we compare our results on the best obtainable magnetic field resolution systematically to other studies of graphene Hall elements in Section 5.7.3.

The decades old, scientific debate on the origin of 1/f noise and the proper theoretical model is continued also for graphene. Different experimental findings support both mobility fluctuations, arising from fluctuations in the scattering cross-section, as well as carrier number fluctuations, arising from trapping-detrapping events due to charged impurities in proximity to graphene, as prevailing microscopic mechanisms of 1/f noise in graphene [389, 392]. Accordingly, the Hooge model and the McWhorter model and combinations of them were discussed to describe 1/f noise in

graphene. Moreover, one needs to distinguish between channel and contact contributions to the 1/f noise in atomically thin materials such as graphene [392, 401]. In order to provide insights in these open questions, resistance noise has been studied for graphene devices with different charge carrier densities, electrostatic environments (substrates), mobilities, layer thicknesses and device areas.

Charge carrier density dependence: Several groups have reported an unusual dependence of the 1/f noise amplitude on the charge carrier density for graphene devices on SiO_2 [389, 392, 401–404]. The noise amplitude was found to have a minimum at the CNP where the resistance is at its maximum. For resistance fluctuations one would not expect such V-shape or M-shape dependence on the gate voltage (density) but instead a Λ-shape dependence where the noise amplitude peaks at the CNP in accordance with the Hooge relation or the McWhorter model [389, 402]. This behavior has been attributed to the spatial charge inhomogeneity near the CNP or the interplay of long-range and short-range disorder [389, 392, 404, 405]. However, for high-κ dielectric substrates and hBN supported or encapsulated graphene, the expected Λ-shape is observed [406–409]. Moreover, the noise voltage between the Hall contacts, i.e. in transverse configuration, exhibits a Λ-shape dependence on the density both for graphene on SiO_2 and hBN-encapsulated graphene (Refs. 395, 397 and the following sections of this work). Takeshita and coworkers (Ref. 402) found that for graphene on SiO_2 the M-shape dependence of the (longitudinal) noise amplitude around the CNP transforms to the expected Λ-shape and Hooge's relation is recovered when the bias voltage is reduced. Their hypothesis is that a (too) large bias voltage induces depinning of electron-hole puddles which results in the anomalous density dependence of the noise amplitude around the CNP. However, this does not explain why the M-shape dependence is not observed for hBN-encapsulated devices. We speculate that this could be related to the relatively larger contact resistance contribution to the total resistance in these devices than in case of graphene on SiO_2 (see impact of contact noise below). Another experiment on dual-gated, hBN-encapsulated graphene allows to separate the contact and channel resistance noise contributions [401]. It shows that long-range charged impurities are the main source of 1/f noise in the channel and channel resistance noise follows Hooge's empirical formula.

Layer dependence: Liu and coworkers (Ref. 410) investigated the atomic layer dependence of the conductance noise from single to > 10 atomic planes for graphene on SiO_2 and found that 1/f noise is a surface phenomenon below seven layers (above seven layers volume noise eventually dominates). That means 1/f noise originates mainly from the surface and varying the film thickness only changes the fraction of current passing through the surface layer [389]. For surface noise, the normalized

noise power spectral density scales with $1/N^2$ (N is the layer number). This corresponds to a reduction of noise amplitude, or Hooge parameter equivalently, by a factor of four comparing monolayer to bilayer graphene which was also reported by another study [411]. For suspended bilayer devices, the reduction of noise amplitude is even stronger and eventually dominated by contact resistance noise [412]. As will be discussed in Section 5.5.3, for our Hall voltage noise measurements in transverse configuration, we observe a reduction of the Hooge parameter by even one order of magnitude comparing hBN-encapsulated monolayer and bilayer devices.

Substrate dependence: The disorder in graphene and its electrical transport properties are strongly influenced by its environment, that means the substrate and encapsulating material. Therefore, it is no surprise that also the 1/f noise level, characterized by Hooge parameter, depends on the choice of substrate. Compared to graphene on SiO_2, suspended graphene, graphene on high-κ dielectric substrates and, in particular, hBN-supported or encapsulated graphene exhibits more than one order of magnitude smaller 1/f noise [404, 406–409, 413]. Moreover, it has been found that the defect density in hBN plays a major role on the 1/f noise level [409]. For our Hall voltage noise measurements in transverse configuration, we observe a similar improvement when encapsulating graphene with hBN, as shown in Section 5.5.3.

Contribution of contact resistance noise: Both channel and contact resistance contribute to 1/f noise, in particular for 2D materials [392, 401, 413]. The lower the 1/f noise in the channel, the more important becomes contact resistance noise. For graphene on SiO_2 the impact of the contacts is smaller [389], while the influence of contact noise is most severe in high quality, hBN-encapsulated or suspended devices [412, 413]. 1/f noise in such high quality devices is found to be governed to a large extent by contact resistance fluctuations. The latter scale with R_c^4 and arise from strong mobility fluctuations in the charge transfer region of the metal-graphene contact because of the fluctuating electrostatic environment [413]. The 1/f noise from contact resistance fluctuations may be described empirically by the Hooge model. The contact resistance noise can be reduced by scaling the device dimensions and the contact area (or the width in case of side contacts in vdW hetereostructures). Moreover, dual-gating with two graphitic gates can minimize the environmental electrostatic fluctuations when one gate is used to screen the hBN-SiO_2 interface that induces more inhomogeneity in the device (due to remote charged impurities) [401, 409, 413]. It was also demonstrated that 1/f noise from the contacts is reduced when the chemical and structural integrity of graphene is preserved under the metal contacts [413]. This could be realized, for instance, by no longer using Cr as adhesion layer for the contacts but only Au as contact material [413, 414]. However, this may result in less reliable contacts from our experience

as Cr serves as an adhesion layer. Alternatively, few-layer graphene could be used to contact the monolayer device, although it was only shown that this reduces the contact resistance but the noise was not studied [415]. To conclude, there is potential room for further improvement of the 1/f noise level in hBN-encapsulated devices when one addresses the contact resistances. However, this also adds complexity in the large-area and mass-scale fabrication of such devices (e.g. few-layer graphene as contacts, dual-gating with graphitic gates). We did not attempt to improve the contact resistances systematically in this work. In all our devices, Cr/Au was used for the electrical contacts. Moreover, in the model used in this work, the contact resistance is simply considered as a parameter when calculating the total resistance (Eq. 5.11), i.e. without taking the exact microscopic mechanism into account. The total resistance is then used in the formula for thermal noise (Eq. 5.10) and the Hooge model for 1/f noise in case of constant bias current operation (Eq. 5.14), as discussed in the preceding section.

Temperature dependence: Several studies have so far addressed the temperature dependence of 1/f noise in graphene devices [392, 404, 409, 411, 416, 417]. In the literature, temperate dependent data is available only for longitudinal resistance fluctuations between cryogenic and room temperature but not for Hall voltage fluctuations in transverse configuration. Unfortunately, the reported results are not consistent. Over a wide temperature range and for devices with lower quality, e.g. on SiO_2, the noise amplitude decreases when the temperature is lowered. However, the extent of this decrease varies among different studies. In some cases, the M-shape dependence of the noise amplitude on the density at room temperature changes to a Λ-shape dependence at lower temperature [404, 416, 417]. In other cases, the Λ-shape dependence is observed at all temperatures [409, 411]. For hBN-encapsulated devices and away from the CNP, it was shown that the temperature dependence of the noise amplitude can be described by the temperature dependence of several different microscopic scattering mechanisms causing the mobility fluctuations (including phonons and thermal activation of trap states) [409, 417]. Finally, the contact noise does not exhibit a substantial temperature dependence [413]. In Section 5.6.2, we will not analyze our temperature dependent noise measurements in transverse configuration by means of such sophisticated models because no notable temperature dependence of the bias voltage *normalized* noise voltage was observed within the studied temperature range, i.e. from room temperature to 80 °C. The observed increase in the actually measured noise voltage with temperature can be attributed to the increase of the resistance which results in an increase of the bias voltage for constant current operation.

5.4. Measurement setup and program to characterize Hall transducers

In this section, we introduce our measurement setup and program which we use to characterize our graphene Hall elements in terms of sensitivity, magnetic field resolution and all other performance indicators including their temperature dependence (from room temperature to about 80 °C).[16] In the first part, all important components and specifications of the setup will be discussed. In particular, our self-built sample box is introduced which can be mounted directly to the electromagnet system. Moreover, the main instruments, e.g. the spectrum analyzer and the low noise current source, and the circuit diagram used for the electrical measurements are presented. The accuracy and limitations of the noise voltage measurement are then investigated in detail in the second part of this section. This includes also a comparison to other electrical setup configurations, e.g. using a different amplifier or spectrum analyzer. Finally, we briefly summarize our measurement program which allows us to explore the large parameter space by varying the charge density, frequency, current and temperature of the devices.

5.4.1. Components and specifications of the measurement setup

Home-built sample box with built-in heater and amplifier and the electromagnet system

CAD drawings of our sample box and pictures of the sample box mounted between the coils of the electromagnet system are shown in Fig. 5.2. A 24 pin chip carrier can be plugged-in at the center of the printed circuit board (PCB). The 4 mm by 4 mm substrate with our sample is glued and bonded onto such chip carrier. The center part of the sample box forms a sealed chamber which can be put under vacuum. It is also equipped with a resistive heating plate with direct contact to the chip carrier when the latter is plugged-in. We tested several heaters and finally ended up with a ceramic heater from Watlow (Ultramic 600 - CER1-01-00335) to which we attached a Pt100 sensor. It allows us to control the temperature between room temperature and up to 120 °C (effective temperature). This means that non-encapsulated devices can be thermally annealed in vacuum inside the chamber. We also glued another

16 The temperature dependence down to cryogenic temperatures in Chapter 4 was performed in a cryostat system and lock-in amplifiers were used for the electrical measurements, as already discussed in that chapter.

Pt100 on top of a chip carrier to properly determine the temperature offset between the heater-attached temperature sensor and the actual, effective temperature on top of the chip carrier. The delta for each set temperature was reproducible and automatically taken into account in the following (larger deviations of more than a few °C were found only above 70 °C). The temperature read-out is performed by a ITC503 from Oxford instruments (calibrated for a Pt100 sensor). At first, we also used this device to power the heater but found that this induces additional spikes in the frequency range of several hundred kHz, likely due to electromagnetic interference (see, for instance, panel d of Fig. 5.3). We then switched to another, standard voltage source. The heating element is dimensioned such that it is not powerful enough to substantially increase the temperature of the entire sample box or PCB in order to not affect the temperature of the amplifier circuits at a distance of more than 10 cm.

The PCB board is also equipped with two amplifier circuits which are battery powered and exhibit very low noise. For each amplifier, two out of 12 pins (i.e. in total all 24 pins) can be arbitrarily selected as input of the amplifier. This allows high flexibility when bonding our samples on the 24 pin chip carrier. We use the AD8429 instrumental amplifier from Analog Devices which is characterized in detail in the following section. The gain resistor is chosen such that the gain is 100 (corresponding to 40 dB). This means that the noise from the sample, which constitutes the signal in our case, is enhanced inside the sample holder before it is transmitted to the spectrum analyzer via well-shielded triax cables, making it much less susceptible to external noise sources. The sample box has a properly grounded, metallic housing which shields the inside from extrinsic electromagnetic interference. Hence, the large printed circuit board with the attached 24 pin chip carrier and integrated amplifier circuit mimics the on-chip amplification and signal processing ASIC in commercial Hall sensors.

Finally, the sample box is mounted between the coils of an electromagnet system (see pictures in panels b and c of Fig. 5.2). We use a water-cooled, resistive coil system from GMW Associates (model 3470) that is controlled by a bipolar, 4-quadrant power supply (Kepco BOP50-8ML-4886). Since the chamber with the sample and heater at the center of the sample box was designed to be as thin as possible, the gap between the two plane poles with 40 mm diameter amounts to only 23 mm. This results in a very uniform magnetic field at the sample position. The magnetic field strength can be adjusted by varying the applied current through the coils. Based on our geometry, we obtain about 100 mT/A and easily reach $300 - 500$ mT in absolute field values. For our measurement, we typically restrict ourselves to $50 - 100$ mT. The cover of the sample chamber has a slit (see panel c) into which a high-precision magnetic field sensor from LakeShore can be inserted as reference

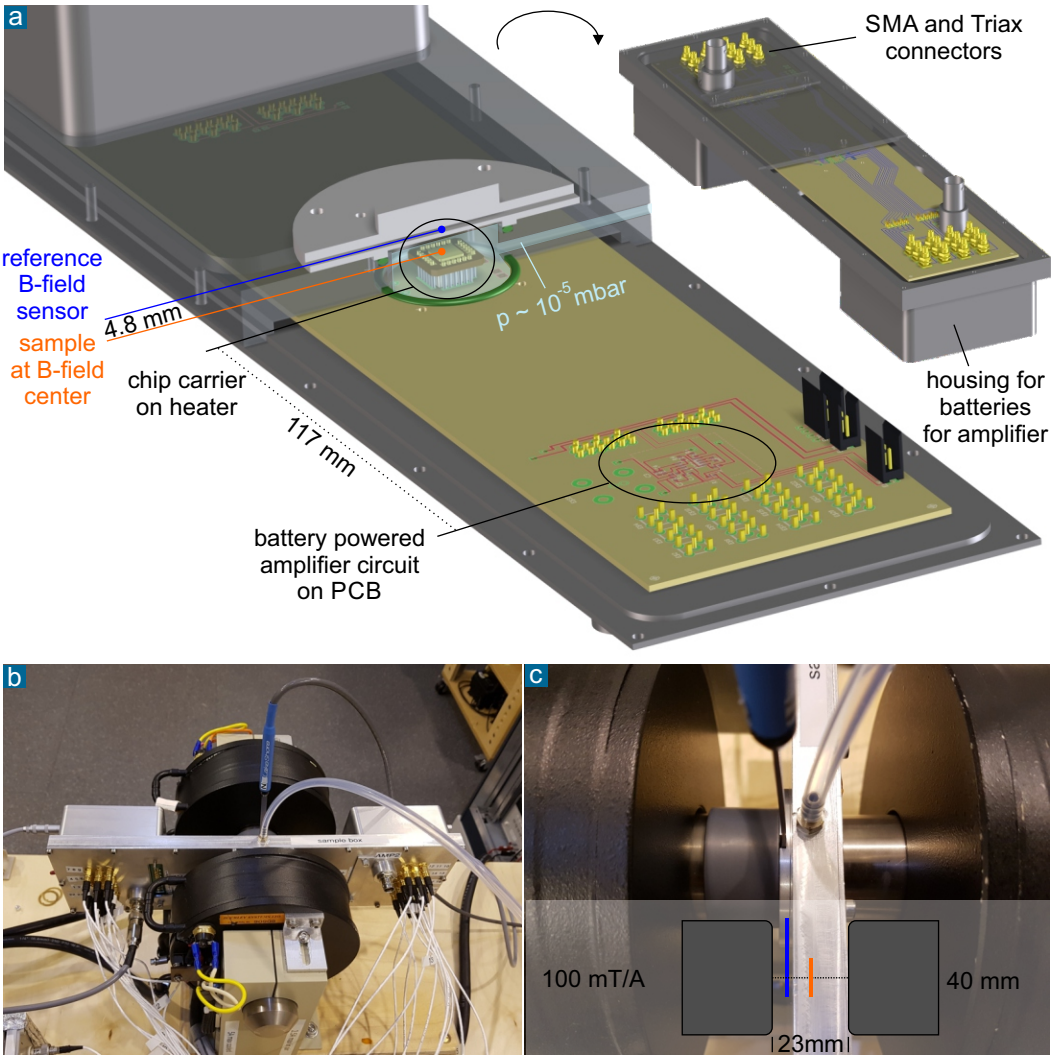

Figure 5.2. a: CAD drawings showing both sides of our home-built sample box. The partial lateral cut through the metallic housing allows a look at the sample chamber and the printed circuit board. The PCB connects the holder with 24 pins in the center, into which the chip carrier with the bonded sample can be plugged in, with two amplifier circuits as well as the electrical SMA and Triax connectors at the outer area of the sample box. Spatially separating the sample holder from the amplifier circuits and the electrical feed-throughs ensures that the sample box can be mounted between the two coils of the electromagnet with an as small as possible gap. This results in a high and homogeneous magnetic field. Moreover, the stray magnetic fields at the position of the amplifier circuits are negligible. Pictures of the mounted sample box are shown in panels b and c.

(Model 475-HMMT-6J04-VR, connected to the corresponding readout device from LakeShore). The distance between our sample and the active area of the reference sensor is only 4.8 mm and their planes are parallel. The difference in magnetic field value is negligible given the very homogeneous magnetic field. Therefore, we use this reference magnetic field sensor to continuously monitor the magnetic field value when sweeping the magnetic field. In general we found the electromagnet system to exhibit only a very small hysteresis. Since we always swept in one direction only and measured the B-field with the reference sensor, this small hysteresis does not induce any errors in our measurements[17].

Setup for the electrical measurement

The whole setup, consisting of the sample box, the electromagnet system and the instruments, was installed in one of the special measurement boxes in the Precision Laboratory of the Max Planck Institute for Solid State Research. This special building as well as these room-sized measurement boxes were designed to provide a ultra low-noise environment. In our case, the electromagnetic shielding from the environment is the most relevant feature. Moreover, each measurement box has its own direct grounding line (not shared with other equipment or experiments). We electrically isolated all measurement instruments from the normal power grid ground through transformers and then connected them to the own grounding line of the measurement box in a star-shaped grounding scheme. Moreover, the GBIP connection to all instruments from the computer running the measurement program was isolated using an optical isolator. These measures notably improved the electromagnetic interference and noise induced by the equipment when measuring the noise of the sample.

The circuit diagram and all used instruments for the electrical measurement of our Hall elements are outlined in panel a of Fig. 5.3. The electrical setup is similar to the setup we used in Chapter 4 where we measured our devices in a cryostat system. In the latter case, the devices were biased by an alternating current and lock-in amplifiers were used to measure the longitudinal and Hall voltages. In contrast, in the setup discussed here a direct current is applied and the voltages are acquired

17 Moreover, the strength of the stray field of the electromagnet declines very fast with the distance from the poles at the center. It does not affect the amplifier circuits located at the edges of the sample box even when they are used while applying high magnetic fields. The stray field at the position of the amplifiers was measured to be about 1/100 of the magnetic field value at the center. This has no impact on the amplifiers. Moreover, this would only have an impact on the Hall coefficient measurement. The noise voltage spectra are recorded in the absence of a magnetic field, as described in Section 5.4.3.

using conventional multimeters (Agilent 34410A). In both cases, the gate voltage is controlled via a source-measurement unit from Yokogawa (SMU GS610) which allows to detect very low leakage current values. In all measurements, one multimeter is used to measure the longitudinal voltage drop between the source and drain contacts of the Hall cross (which allows to calculate the two terminal resistance). In case of a Hall bar geometry, also other contact pairs can be monitored using additional multimeters. Another multimeter is always connected in parallel to the spectrum analyzer to read out the output voltage of the amplifier, i.e. the amplified Hall voltage. The constant direct current is applied between the source and drain contact by means of a high precision, ultra low noise current source from Stanford Research Systems (model CS580). It is equipped with an external GBIP interface that is connected to the current source via an optical interface which again reduces any (external) noise or interference. As demonstrated in Fig. 5.4, the current source is as good as using a 9V block battery and pre-resistor for applying a low noise current, however, offering full control and flexibility.

Finally, the inputs of the above mentioned, low noise instrumentation amplifier (Analog Devices AD8429) are connected to the Hall contacts of the device. The amplifier circuit on the PCB board is designed such that the voltage gain is 100. The gain is constant over a wide frequency range but eventually starts to decrease slightly above 400 kHz (see discussion further below and yellow boxes in panels c and d of Fig. 5.3). We found the amplifier to possess an ultra low noise level, comparable or even better than other external amplifiers, as well as to provide a very clean noise voltage spectrum of our samples, i.e. the noise voltage between the Hall contacts. This will be discussed in more detail in the following Section 5.4.2 and is demonstrated in Fig. 5.4. This detailed analysis will also include the other key instrument of our noise voltage measurement setup, the spectrum analyzer. We have chosen a swept-tuned, super-heterodyne spectrum analyzer (Agilent E4440A) which is operated in DC coupling mode in order to be able to measure also the kHz and MHz frequency range of interest. The instrument received maintenance and a new calibration by an OEM certified service partner before it was installed in our setup. Its noise floor is determined and discussed in the following section. In general, the noise floor is larger than for FFT based spectrum analyzers. However, super-heterodyne spectrum analyzers are able to measure higher frequencies than the conventional FFT analyzer (typically limited to around 100 kHz). Since we want to acquire spectra up to 1 MHz, which allows us to properly determine the corner frequency and the transition from 1/f to thermal noise also in case of our few micrometer sized devices, the Agilent E4440A is the preferred solution for us. Another drawback of the measurement principle of super-heterodyne spectrum analyzers is that their intermediate frequency mixing stage (IF stage) is sensitive to high DC

voltages. Since we have to use the device in DC coupling mode, we must pay attention that we do not apply too high DC voltages (should be below 50 mV). Our noise measurements are performed in the absence of a magnetic field, i.e. the Hall voltage is zero. Nonetheless, the Hall offset can result in a notable DC voltage when larger bias currents are applied. The offset voltage is further enhanced by the amplifier. Let us assume that the Hall offset resistance has a maximum of around 50 Ω and we apply a current of 30 μA, then the offset voltage amounts to 1.5 mV. Given a gain of 100, the applied DC voltage to the spectrum analyzer is 150 mV which would overload the IF stage and not allow to perform a measurement. Therefore, we either have to set a suitable input attenuation, which is possible to do in the spectrum analyzer settings, or use an external high pass element to filter the DC part. Both solutions may have different consequences on the measured noise spectra that one needs to consider as will be discussed in more detail in the following. We found that it was usually enough to set an attenuation of 10 dB which has no notable impact on the obtained spectra.

5.4.2. Accuracy and limitations of the noise voltage measurement using our electrical setup

Analysis of the noise voltage measurement process

As already discussed in Section 5.2.2, different noise sources add in root sum of squares manner to the total noise voltage spectral density when they are independent and uncorrelated (see Eq. 5.6). According to the preceding theory section we expect that thermal and 1/f noise will be the prevailing noise sources in our graphene devices (see $N_{V,therm}(R)$ in Eq. 5.10 and $N_{V,1/f}$ in Eq. 5.14 in Section 5.3.1). Our goal is to determine their combined noise voltage spectrum as accurate as possible. Our metallic sample box and the star-shaped grounding scheme of all electrical instruments suppresses external noise, e.g. EMI, reliably. Moreover, the signal of interest, in our case the noise level of the graphene Hall element, is amplified with a gain factor of 100 (+40 dB), g_{amp}, before it is transmitted to the spectrum analyzer. Due to the peculiarity of the DC coupling stage of the spectrum analyzer and the output resistance of the amplifier circuit, they act as a voltage divider (-6 dB), as discussed above. Accordingly the effective gain factor, g_{eff}, is 50 (+34 dB). We also confirmed this behavior by applying DC and AC voltages of different frequencies directly to the amplifier input and measuring the value with the spectrum analyzer.

Like any electronic component, the amplifier has intrinsic noise. In the specification sheet and manual of the amplifier (AD8429 from Analog Devices), formulas are provided to calculate the noise of the input stage of the amplifier. It consists of three parts: The voltage noise, $N_{V,amp}$, and the current noise, $N_{I,amp}$, of the instrumentation amplifier as well as the thermal noise of the source resistance connected to the input stage. The latter is mostly given by the thermal noise of the two terminal resistance of our sample since the other resistors used in our amplifier circuit at the input stage are substantially smaller and hence negligible (2x 33 Ω). Therefore, we focus on the voltage noise of the amplifier which amounts to $1.5\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ at the input stage for a gain of 100 as well the current noise of $1.5\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$. The noise voltage spectrum density of the latter is then given by multiplication with the source resistance at the input stage which is essentially the sample resistance, as discussed. Finally, the 1/f noise corner frequency of the amplifier is about 10 Hz such that we can neglect 1/f noise from the amplifier in the following.

To summarize, the input to the spectrum analyzer (SA), that means the combined noise voltage spectrum of our device under test and the amplifier, after amplification and passing the voltage divider, is given by

$$N_{V,SA-IN} = \sqrt{N_{V,therm}^2(R) + N_{V,1/f}^2(I > 0) + N_{V,amp}^2 + N_{I,amp}^2(R)} \cdot g_{eff} \ , \quad (5.21)$$

where R is essentially the resistance of our sample, i.e. the resistance between the Hall contacts, and I is the applied current. In case no current is applied, the 1/f noise term can be neglected. Apart from the graphene devices, we also connected different resistors to the amplifier inputs for further testing, as discussed below. Then, R corresponds to this resistor value.

The spectrum analyzer allows to set different attenuation levels (0 to 20 dB) which translates to corresponding attenuation factors, $a_{SA} < 1$. As discussed above, it is sometimes necessary to set such attenuation to protect the IF stage of the spectrum analyzer from overloads due to a too high DC voltage (resulting from the Hall voltage offset at larger bias currents). The attenuated signal is then mixed with the frequency dependent noise floor of the spectrum analyzer, $N_{SA-NF}(f)$. Subsequently, the mixed signal is automatically multiplied by the spectrum analyzer software with a factor that compensates for the attenuation, i.e. by $1/a_{SA}$. The spectrum analyzer output is then given by

$$N_{V,SA-OUT} = \sqrt{(N_{V,SA-IN} \cdot a_{SA})^2 + N^2_{SA-NF}(f)} \cdot (1/a_{SA}) \ . \tag{5.22}$$

This corresponds to the raw noise voltage spectrum which we read out from the spectrum analyzer. Finally, we obtain the noise voltage spectrum of our sample, for which we will use the notation $N_{V,meas}$ in the following (Eq. 5.24 in Section 5.5), by simply dividing the spectrum analyzer output by the effective gain factor:

$$N_{V,meas} = N_{V,SA-OUT} \, / g_{eff} \ . \tag{5.23}$$

Potential sources of inaccuracy and error estimation

By using Eq. 5.23, we do not differentiate between the noise of the sample and the amplifier. This would be in principle possible by applying Eq. 5.21. However, as discussed below, the induced error by this simplification is negligible. If at all, it only leads to a small inaccuracy when determining the thermal noise floor in case that the two terminal resistance of the device is higher, above 5 kΩ, which is typically not the case. Moreover, thermal noise dominates only at higher frequencies in case of our few micrometer sized devices (see following section). Our main interest is the determination of the 1/f noise which prevails in the low to medium frequency regime as soon as just a small current is applied.

If not taken into account properly, a severe source of inaccuracy can be a too high attenuation level at the spectrum analyzer's input, in particular for the frequency range below 20 kHz. In panel c of Fig. 5.3, the noise floor of the spectrum analyzer, $N_{SA-NF}(f)$, is given for different attenuation levels. It was obtained without connecting any other equipment to the spectrum analyzer input. The green curve (0 dB) corresponds to the real, physical noise floor while the other two curves (10 dB and 20 dB) correspond to the effective noise spectra after mathematical compensation of the attenuation by the spectrum analyzer software, i.e. after multiplication of the physical noise floor by $1/a_{SA}$. We then re-connected the amplifier with a 2 kΩ resistor between its inputs to the spectrum analyzer and measured the noise voltage spectrum without attenuation and with 20 dB attenuation (bluish curves in panel c). We find the spectrum to be flat from 300 Hz onwards when no attenuation is set (dark blue data). Below 300 Hz we observe a slight impact of the spectrum analyzer

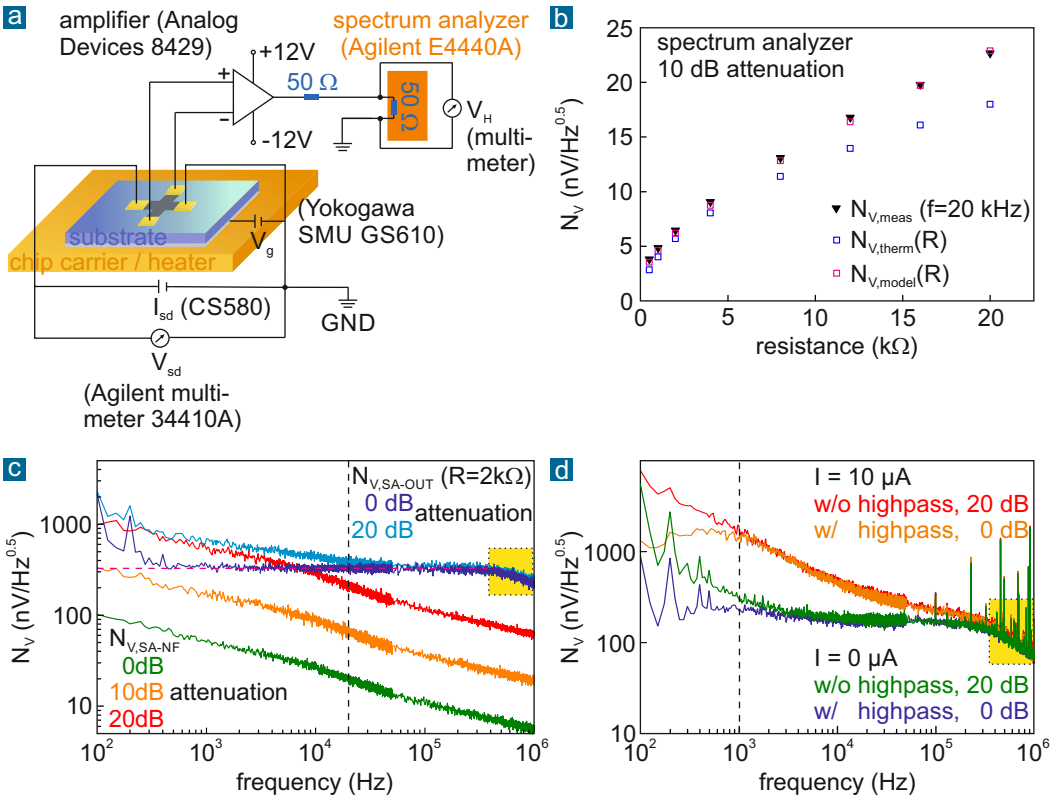

Figure 5.3. a: Instruments and circuit diagram used for the electrical measurement (see main text). b: Measured noise voltage of different resistors between the amplifier inputs at 20 kHz and 10 dB attenuation (black triangle). The measured data are compared to the calculated values considering only the thermal noise of the resistor at room temperature using Eq. 5.10 (blue squares) as well as when taking into account the amplifier noise and the noise floor of the spectrum analyzer, i.e. our model based on Eqs. 5.21 and 5.22 (magenta squares). c: Effective noise floor of the spectrum analyzer, $N_{V,SA}$, for different attenuation levels (0, 10 and 20 dB) and the obtained noise voltage spectrum, $N_{V,SA-OUT}$, of a 2 kΩ resistor between the amplifier inputs in case of 0 and 20 dB attenuation. The yellow box highlights the onset of the frequency dependent gain of the amplifier. d: Noise voltage spectra of a graphene device with and without applied current. Configurations with high pass and 0 dB attenuation as well as without high pass but with 20 dB attenuation are compared, as discussed in the main text.

noise floor which is in addition overlaid by extrinsic spikes at 50 Hz multiples. Moreover, the noise voltage starts to decrease above 400 kHz (yellow boxes in panels c and d). This is because of the frequency dependent gain of the amplifier above 400 kHz, in accordance to its specification sheet.

In case of 20 dB attenuation (turquoise data), the noise voltage of the 2 kΩ resistor and the amplifier become too small and N_{SA-NF} dominates at lower frequencies (below 20 kHz). This may cause an incorrect interpretation of the measured spectrum of the resistor (frequency dependent instead of flat) and, accordingly, an inaccurate estimate of the noise voltage values of the device under test. However, this notably affects only the measurements of the noise floor of our samples, i.e. when no bias current is applied, and only at lower frequencies (below 20 kHz). The 1/f noise in our devices is typically at least one order of magnitude larger than this noise floor, even for small bias currents, and the corner frequencies are in the range of several tens of kHz. An example of such a 1/f noise spectrum in case of a biased graphene Hall element is given in panel d of Fig. 5.3. While the use of 20 dB attenuation has an impact on the measured noise floor below several kHz (compare green to blue curve), the 1/f noise dominated spectrum of the biased device is not affected (red curve). In the following, we will extract the 1/f noise voltage spectrum, $N_{V,1/f}$, from our raw data by subtracting the measured noise voltage spectrum in the absence of a bias current from the spectrum with applied current (see Eq. 5.24 in Section 5.5.1). Even in that case, the induced error is marginal since the noise voltage at lower frequencies is significantly larger when a current is applied in comparison to the noise floor without applied current (at least for frequencies above 500-1000 Hz). Therefore, the relatively small change to the noise floor due to the attenuation is negligible. Finally, we always kept the attenuation as low as possible. Typically, 10 dB was sufficient. This already affects the noise floor measurement much less than in the extreme case of the previously discussed 20 dB attenuation.

As an alternative to attenuation, we also tested the possibility of inserting a high pass filter with a cut-off frequency at about 100 Hz between the amplifier output and the spectrum analyzer input in order to eliminate the DC offset voltage. This allowed performing all measurements without attenuation. However, we found that the noise voltage spectrum is then distorted up to 1 kHz by the high pass filter[18], as shown in panel d of Fig. 5.3. Since this also affects the noise voltage spectrum of the biased device, i.e. the 1/f noise spectrum of main interest here, we decided to perform the measurements without such a high-pass filter. The use of a reasonable attenuation level is the superior approach instead, since it does not notably affect the accuracy of our measurement.

Let us now come back to the above test case of the 2 kΩ resistor connected to the amplifier inputs. As shown in panel c of Fig. 5.3, a noise voltage floor $N_{V,SA-OUT}$ of

18 The spikes in the frequency range of several 100 kHz arise from electromagnetic interference picked up or induced by the temperature controller connected to the ceramic heater (Oxford ITC503). We later used the ITC503 only for the read-out and connected another voltage source to power the heating element. This removed these spikes almost completely.

about $310\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ can be determined from the data at 0 dB attenuation (see dashed magenta line). This results in $N_{V,meas}$ of $\approx 6\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ (Eq. 5.23). The obtained value is very close to the expected thermal noise floor for a $2\,\mathrm{k\Omega}$ resistor at room temperature, $N_{V,therm} \approx 5.7\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ (see Eq. 5.10 in Section 5.3.1). The error is about 5 %. We repeated the measurement for several different resistor values at 20 kHz with 10 dB attenuation. These parameters correspond closely to those later used when measuring our graphene Hall elements (the corner frequency of 1/f noise is typically above 10 kHz even for the smallest bias currents). The $N_{V,meas}$ values (black triangles) are shown in panel b of Fig. 5.3. They are compared to the theoretical $N_{V,therm}$ values for each resistor (blue squares) as well as to the calculation using the full model discussed above (Eqs. 5.21 and 5.22), considering both the voltage and current noise of the amplifier. The measurement and the model fit nicely. Below $5\,\mathrm{k\Omega}$, the difference between the model (or the measured data) and the theoretical thermal noise floor of the resistor is marginal. For larger resistances, the deviation increases because of the current noise of the amplifier which scales with the source resistance, as discussed above. The two-terminal resistance in our graphene devices typically did not exceed $5\,\mathrm{k\Omega}$ (at the CNP). Hence, the assumption in the following section that $N_{V,therm} \approx N_{V,meas}(I = 0)$ is justified, in particular since $N_{V,meas}(I = 0)$ is only used to extract the 1/f noise from the $N_{V,meas}(I > 0)$ data and we do not intend to determine the precise resistance value (which would be possible, when using the full model).

Finally, the 1/f noise in our few micrometer sized devices is at least several tens of $\mathrm{nV}/\sqrt{\mathrm{Hz}}$ at 1 kHz even for low current values. For instance, $60\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ as shown in Fig. 5.6 in the following Section 5.5.1. Accordingly, $N_{V,1/f}$ clearly dominates $N_{V,SA-IN}$ which amounts to about $3000\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ in case of an effective gain of 50. For the frequently used attenuation of 10 dB, one eventually obtains approximately $1000\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ which is substantially larger than the physical noise floor of the spectrum analyzer, $50\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ at 1 kHz. The 1/f noise contribution is also about one order of magnitude larger than the noise floor of our sample (and amplifier). Accordingly, the error induced by the small increase of the actually measured noise floor due to the spectrum analyzer noise floor at 10 dB attenuation is small. All in all, we are confident about the accuracy of our data, in particular above 1 kHz. We estimate the maximum error to be typically a few percent only.

Comparison to other electrical setups

In order to demonstrate that our electrical setup is capable to reliably and accurately acquire very clean noise voltage spectra of Hall elements, we measured one

and the same sample using different setups. This comparison is shown in Fig. 5.4. At first, our standard configuration involving the CS580 from Stanford Research as current source, the sample box integrated and battery powered amplifier AD 8429 from Analog Devices as well as the Agilent E4440A spectrum analyzer (10 dB attenuation) is used to obtain the noise voltage spectrum both without applied current (black curve) and with applied current (green curve). This corresponds to configuration 1 in panel b. In configuration 2, the integrated amplifier was replaced by a conventional, external amplifier, SR560 from Stanford Research, which was also battery powered (orange curve). The comparison to our standard configuration demonstrates that much more extrinsic interference is picked up and the obtained spectrum is less clean and accurate. This is likely because the signal from our sample, i.e. its noise, is transmitted to the external amplifier at a lower amplitude and over a larger distance using a conventional BNC cable. In configuration 3 (purple data), we replaced the low noise current source, the CS580, with a 9V block battery and pre-resistor to apply the same current of about $30\,\mu$A as in the other cases. Both spectra are quasi identical, apart from some external noise spikes in the spectrum. This demonstrates that the programmable current source CS580 is as good as a battery.

Finally, we replaced our super-heterodyne spectrum analyzer, the Agilent E4440A, with a FFT signal analyzer in configuration 4. Again we chose a commonly employed device, the SR780 from Stanford Research. However, its bandwidth is limited to only 100 kHz (note that being able to measure up to 1 MHz in order to determine the corner frequency even in case of high 1/f noise amplitude was one of our criteria why we chose the Agilent spectrum analyzer). Since also the number of data points is limited and we wanted to have a reasonable resolution bandwidth, i.e. frequency steps, we restricted the measured frequency range to 50 kHz. This is also the reason why the frequency axis in panel a of Fig. 5.4 is limited. We manually acquired several data points at different frequencies both without and with applied current (red squares and triangles, respectively). They agree nicely with the spectra obtained with our standard setup (configuration 1)[19]. However, when looking at the whole spectrum (not shown), we found it to be less clean and more noisy than the spectrum we obtain from the Agilent E4440A. Finally, another advantage of the SR780 is that its internal noise floor is very low, around $5\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$. This allows us to connect it directly to the Hall voltage contacts of our device, i.e. without the need for an amplifier (configuration 5). Of course, the obtained spectra were even more noisy but we could extract, again manually, several data points after some time averaging

19 We note that the SR780 does not act as a voltage divider in combination with our amplifier circuit such that the gain factor was 100 effectively (which was considered for the plotted data, see also last column of panel b).

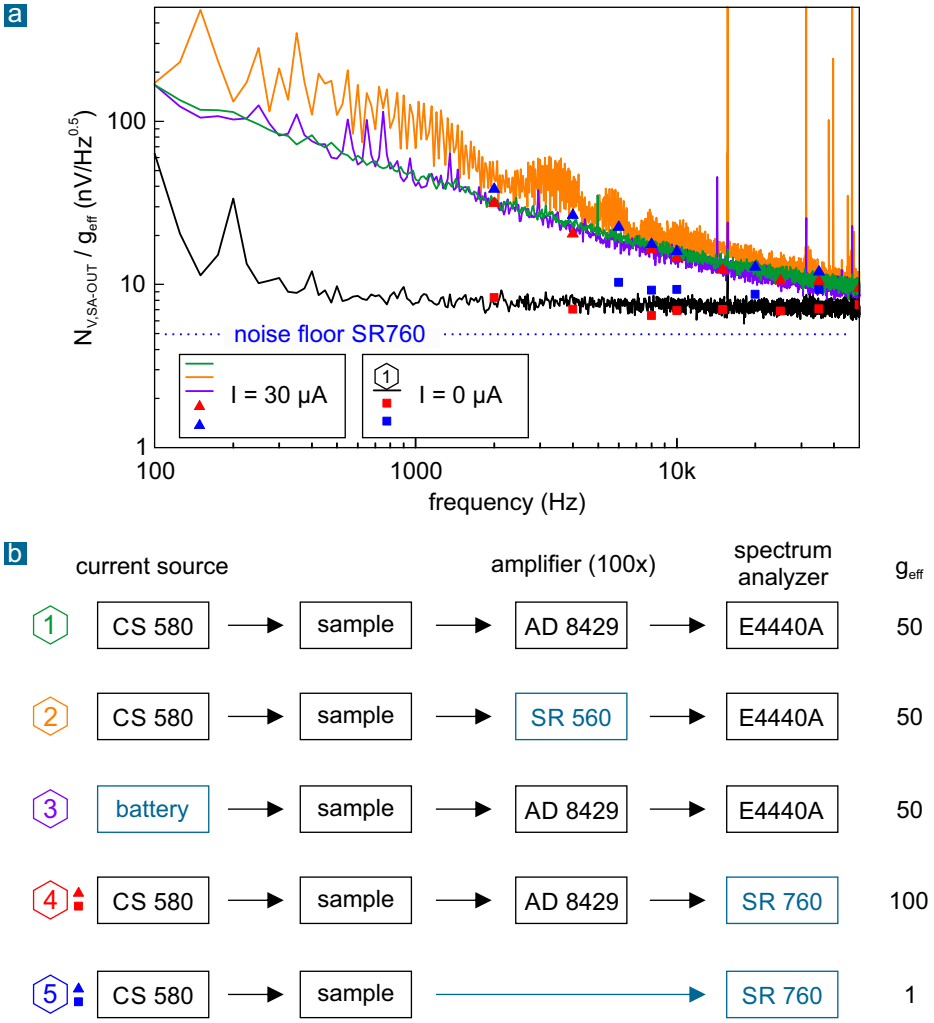

Figure 5.4. Comparison of different electrical setups for noise voltage measurements on the same graphene Hall element (both with and without applying a current). Various combinations of current sources, amplifiers and spectrum analyzers are tested. The data in panel a is colored according to the configuration number in panel b. The thermal noise floor (i.e. without current) is measured only for the first, fourth and fifth configuration. See main text for details.

(blue squares and triangles, respectively). Again, these acquired data points coincide reasonably with the spectra that are obtained using our standard configuration (in particular, when we take into account that the signal, i.e. the noise voltage of the sample, is of the same order of magnitude than the noise floor of the SR780

spectrum analyzer). Since configuration 5 does not involve any amplification step and the gain factor is one, its agreement with the data obtained by our standard setup configuration confirms again that the effective gain is really 50 (and not 100) when combining the amplifier circuit and the Agilent spectrum analyzer.

To summarize, the comparison of the different electrical setups demonstrates that the noise voltage measurement with our standard configuration is reliable, accurate and results in the cleanest spectrum. This is not always the case for other noise spectra reported in the literature which are often less clean, more noisy or exhibit a large amount of spikes due to electromagnetic interference (e.g. Refs. 198, 348, 349). Other studies also managed to achieve quite clean spectra but typically did not acquire noise spectra beyond 1 or 10 kHz (e.g. Refs. 397, 398, 418, at which Ref. 418 measured up to 100 kHz). Our measurement up to 1 MHz allows us to properly determine the crossover between 1/f and thermal noise.

5.4.3. Measurement program

In the following, we briefly describe our measurement program that performs all steps fully automated. The detailed measurement of one device takes at least one day (depending on the number of gate voltage steps) and involves the acquisition of both the Hall coefficient and the noise voltage spectra from 100 Hz to 1 MHz for many different gate voltage steps as well as for various bias currents. If several temperature steps are desired, the measurement time scales accordingly (up to one week).

At first, the measurement is prepared by checking all electrical contacts of the sample work properly and the contact resistances are not unusually high. In case of the Hall cross geometry, we also confirmed that the resistances of both contact pairs are similar, i.e. that the device is rotation-symmetric. Then, a small current is applied, e.g. 1 μA, and the gate voltage is swept while recording the two-terminal resistance. Moreover, we obtain the Hall voltage curves from a coarse magnetic field sweep for each gate voltage step. From this quick measurement, many relevant parameters such as the resistivity, the Hall coefficient as well as the Hall voltage offset (at zero magnetic field) can be obtained as a function of gate voltage. This allows determining V_{CNP} and adapting the studied gate voltage range accordingly (symmetric around V_{CNP}). Finally, we can estimate, from the Hall offset resistance, the maximum current values that can be applied in order to not exceed the DC voltage input limit of the spectrum analyzer (below 50 mV). If necessary, the attenuation level is

adapted to allow at least current values up to $20 - 30\,\mu$A. Such current values are necessary to properly determine the 1/f noise contribution.

The main measurement program is structured as follows:

1) The temperature is set and an appropriate waiting time ensures that the temperature stabilizes.

2) The gate voltage is swept from the minimum to the maximum value in predefined steps.

3) For each gate voltage step the following measurements are performed sequentially:

   (i) A small current, e.g. $1\,\mu$A, is applied and the magnetic field is swept from about -100 to +100 mT by controlling the current applied to the electromagnet coils. The exact magnetic field value is monitored continuously through the reference magnetic field sensor. The longitudinal voltage drop between the source and drain contacts, which allows to determine the two-terminal resistance, and the output of the amplifier which is connected to the Hall contacts are acquired in parallel. If applicable, e.g. in case of Hall bar devices, also other voltages are measured.

   (ii) The magnetic field is turned off.

   (iii) The noise voltage spectrum between the Hall contacts is measured when no current is applied as well as for at least two or three bias current values. As discussed above, the amplifier is used to enhance the noise from our sample. We used a maximum current density of about $10\,\mu\text{A}/\mu\text{m}$. This means that we typically applied $0\,\mu$A, $10 - 20\,\mu$A, $20 - 30\,\mu$A as well as $30 - 50\,\mu$A in case of our few micrometer sized devices. For each current value, the noise voltage spectrum is acquired in two parts in order to have sufficiently dense data points while keeping the measurement time reasonable. First, the 0 to 50 kHz frequency range is measured using a resolution band width (RBW) of 20 Hz and 2001 data points, i.e. 25 Hz increments. This takes about 90s per measurement. Then, the full 0 to 1 MHz frequency range is measured using a RBW of 100 Hz and 1001 data points, i.e. 1 kHz increments. This takes about 200s per measurement. The overlap of the two data sets (0 to 50 kHz) allows to test whether the two measurements coincide. Since this was always the case, the first 50

data points of the second measurement with 1 kHz steps are disregarded and the two data sets are combined before plotting and further analysis (see for instance Fig. 5.6).

4) The next temperature is set and the same measurement sequence is repeated.

We note that we determine the Hall coefficient from Hall voltage curves, $U_H(B)$, that are obtained using a small current, e.g. 1 μA. We do this in order to protect the spectrum analyzer input which is kept connected such that the above measurement program can be run fully automated without the need to manually plug in or out certain devices. However, the noise voltage spectra are obtained at larger currents. In Section 5.6, we then determine the magnetic field resolution using the measured noise voltage and the corresponding Hall voltage given by multiplying R_H and the current value for which the noise voltage was measured (see definition of B_{min} in Eq. 5.27). In case of (very) large current values, the Hall coefficient curve (as a function of gate voltage or density) may be altered when the potential drop along the sample becomes too large. In separate measurements on several devices, we acquired the Hall coefficient curves for several different current values (e.g. for the same current values used for the noise measurements). We never observed a notable difference in the Hall coefficient curves. Hence, up to the maximum current density we restricted us to, i.e. 10 μA/μm, the Hall coefficient does not depend on the current, as it should. Hence, the Hall voltage can always be calculated by $R_H \cdot I$, as presumed in Eq. 5.27. Furthermore, we note that the noise voltage spectra are measured without applying a magnetic field. This is common practice when characterizing Hall elements and is done in the same way by all other studies in the field (for instance, Refs.348, 349, 398 etc.). We also acquired some noise voltage spectra while applying magnetic fields of up to 100 mT and observed no notable difference to the measurement without magnetic field (data not shown).

5.5. Noise voltage in graphene Hall transducers

In this section we will study voltage fluctuations between the Hall contacts, i.e. the noise voltage spectral density in transverse configuration, for a multitude of graphene Hall transducers. As discussed in the preceding section, our setup allows for precise noise voltage measurements with $nV/\sqrt{Hz}$ resolution. The noise voltage is typically obtained in the absence of a magnetic field while the device is biased

with a constant DC current[20]. The noise voltage spectrum is acquired from 100 Hz to 1 MHz while the charge carrier density is varied at least in the range of $\pm 10^{12}\,\mathrm{cm}^{-2}$. The noise measurement is performed both without applying a bias current as well as for several different currents in order to distinguish between the white noise floor and 1/f noise. This allows us to properly identify all dependencies of the noise voltage on the operating parameters in graphene Hall elements. In the following discussion, we will also include experimental findings on graphene Hall transducers that were reported in the literature. However, there are fewer studies of the noise voltage in transverse than in longitudinal configuration. In particular, often no systematic investigation of the entire parameter space (density, frequency and current) is conducted. Most publications focus on the magnetic field resolution, i.e. the key Hall transducer performance indicator, and do not provide extensive data on the noise voltage alone. In Section 5.7.3, we will compare our best obtainable magnetic field resolution values to these publications on graphene Hall elements (for instance, 198, 348, 349, 397–400).

In the first subsection, 5.5.1, the noise voltage spectral density between the Hall contacts is provided as a function of gate voltage, i.e. density, and frequency. By comparing the measurements without bias current and for several different current values, the interplay between 1/f noise and the white noise floor is demonstrated. The corner frequency is introduced marking the crossover from 1/f noise in the low frequency regime to the white noise floor that is found to be governed by thermal noise. In the second subsection, 5.5.2, it is proven that the 1/f noise between the Hall contacts, i.e. in transverse configuration, can be described by the Hooge relation. All dependencies in this model will be verified. The frequency, current and areal dependence are demonstrated directly, while the density dependence is found to be more complex due to the additional density dependence of the two terminal resistance in case of constant current operation in our measurements. Therefore, the density dependence will be shown indirectly by extracting the Hooge parameter, α_H, from the bias voltage normalized noise voltage spectral density. This is done in the third subsection, 5.5.3, for a large number of different devices. Graphene devices on SiO_2 and hBN-encapsulated graphene devices, i.e. samples with notably different quality and charge carrier mobility, are compared. Moreover, we distinguish between monolayer and bilayer graphene. Finally, the temperature dependence of the noise voltage will be discussed at a later point in this chapter, in Section 5.6.2, when noise

20 Note that we also conducted measurements on biased devices while applying magnetic fields of up to 100 mT. No notable difference in the noise spectrum (from 100 Hz onwards) was observed. However, as discussed in Section 5.4, the spectrum analyzer input is very sensitive regarding a too high DC voltage level which is easily reached through the Hall voltage arising from such magnetic field. Furthermore, the measurement of the noise voltage spectrum in the absence of a magnetic field is a common procedure applied in the literature to characterize Hall elements.

voltage and Hall coefficient measurements are combined to determine the magnetic field resolution.

5.5.1. Interplay of thermal and 1/f noise

Our measurements of the noise voltage spectral density as a function of gate voltage and frequency are summarized in the colormaps of Fig. 5.5. For each gate voltage step, the noise voltage spectrum is recorded from 100 Hz to 1 MHz. The data are acquired both in the absence of a bias current as well as for various different DC current values. This allows us to explore the entire operating parameter space. We present here and in the following mainly data which were obtained on one exemplary hBN-encapsulated graphene Hall element. However, the measurement and analysis were repeated for several other devices with similar results.

Panel a of Fig. 5.5 shows the white noise floor which obviously is slightly higher around the CNP (here, $V_{CNP} \approx 0.5\,V$). This means that the noise floor scales with the two terminal resistance between the Hall contacts. The latter peaks at the CNP. As already shown in the previous Section 5.4, the white noise floor of our setup is mainly determined by the thermal noise of the resistance between the amplifier inputs. That means in our case the two terminal resistance between Hall voltage contacts, i.e. $N_{V,meas}(I = 0) \approx N_{V,therm}(R)$. For the device studied here, we determined the two terminal resistance, including contact resistance, to be about $4\,k\Omega$ at the CNP. This corresponds to a thermal noise of $8\,nV/\sqrt{Hz}$ at room temperature (see Eq. 5.10). The actually measured noise voltage at the CNP is about $10\,nV/\sqrt{Hz}$. Aside from the dominating thermal noise of the Hall element, the second largest noise contribution stems from the amplifier noise (see Eq. 5.6 in Section 5.2.2 as well as the discussion in Section 5.4). Finally, the horizontal lines in the colormaps, in particular at very low frequency and in the several hundred kHz range, correspond to spikes in the noise spectrum due to external noise sources, e.g. electromagnetic interference at multiples of 50 Hz as well as at several hundred kHz. These spikes are also visible in the frequency spectrum for fixed gate voltage in Fig. 5.6. The setup was continuously optimized and most of these external noise sources could be reduced after the data shown here were acquired[21]. However, since also the majority of the presented Hall coefficient data in the previous chapter was obtained on this device, we decided to present the data obtained on the same device in order to have

21 For instance, the grounding scheme of all electrical devices was further optimized and the power controller of the heating stage was replaced which almost completely eliminated the EMI at several hundred kHz.

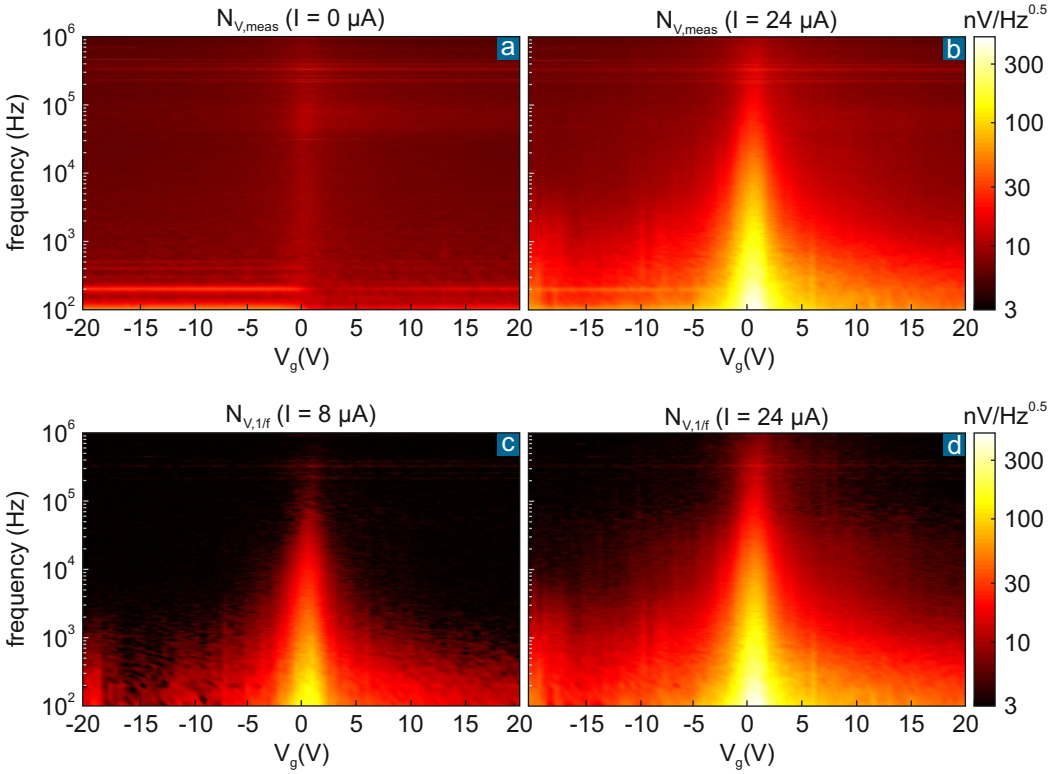

Figure 5.5. Noise voltage spectral density between the Hall contacts as function of gate voltage and frequency for a hBN-encapsulated device (in the absence of a magnetic field). Panels a,b show the measured noise voltage at zero and finite current, respectively. From these original data the 1/f noise voltage contribution is obtained using Eq. 5.24. Panels c,d provide results for a bias current of $8\,\mu$A and $24\,\mu$A, respectively. The boundary between the black and red data points corresponds to the corner frequency. It depends on the gate voltage (i.e. density), as discussed in the main text. The corner frequency shifts to higher frequencies as the current is increased.

a consistent data set. This is in particular important when we later determine the magnetic field resolution.

Panel b of Fig. 5.5 shows the measured noise voltage for the highest current that we applied to the device. Comparing panels a and b, one can clearly distinguish the crossover between 1/f noise and the thermal noise floor which we already discussed theoretically in Sections 5.2.2, 5.3.1 and 5.3.2. The 1/f noise can be extracted using Eq. 5.6 in Section 5.2.2 where $N_{V,tot} = N_{V,meas}$ and when assuming that 1/f noise is the prevailing noise type that scales with the current:

$$N_{V,1/f}(I > 0) = \sqrt{|N_{V,meas}(I > 0)^2 - N_{V,meas}(I = 0)^2|}\ . \tag{5.24}$$

The results are shown panels c and d of Fig. 5.5 for two different current values. As will be discussed in more detail in the following subsections, 1/f noise has a maximum at the CNP and then decreases towards higher density. Moreover, it decreases with frequency and increases with current, as can be seen from panels c ($I = 8\,\mu A$) and d ($I = 24\,\mu A$). The transition from 1/f to thermal noise occurs in these $N_{V,1/f}$ plots at the crossover from red to black data points that mark the corner frequency, defined as the frequency value for which $N_{V,1/f}(f_c) = N_{V,therm}$ (see Section 5.3.1). The corner frequency is a function of both density, i.e. gate voltage, as well as current, This is again nicely demonstrated in panels c and d.

The interplay between 1/f and thermal noise becomes even more evident when plotting the noise voltage data versus frequency for a fixed gate voltage value. This corresponds to a vertical line cut through the colormaps in Fig. 5.5. An example for the largest 1/f noise level at the charge neutrality point ($V_{CNP} \approx 0.5\,V$) is shown in Fig. 5.6. Panel a displays the measured noise voltage spectral density for three different current values and when no current is applied. The extrinsic origin of the spikes at several hundred kHz was already discussed above. The dashed magenta line marks the $1/f^{0.5}$ dependence predicted by the Hooge model. In the low frequency regime, this trend line nicely fits to our data. The corner frequency, f_c, can be graphically determined as the frequency value for which this $f^{0.5}$-line and the noise floor, when no current is applied, intersect.

Panel b of Fig. 5.6 shows the extracted 1/f noise using Eq. 5.24. Apart from the strongly fluctuating data points in the high frequency regime, when the 1/f noise approximately equals the thermal noise floor, the $1/f^{0.5}$ trend line (dashed magenta line) describes the data properly over the entire frequency regime. An alternative plot of the data in panel c by multiplying with $f^{0.5}$ brings this out even better. The result is nearly independent from the frequency. Finally, panel d shows the 1/f noise voltage divided by the bias current. The almost perfect overlap of all curves demonstrates nicely the linear current dependence in accordance with the Hooge model (Eq. 5.14 in Section 5.3.1).

5.5.2. Applicability of the Hooge model for 1/f noise

Our data obtained for a large operating parameter space, i.e. frequency, density and current, allow us to validate the applicability of the Hooge model to describe the 1/f noise between the Hall voltage contacts, i.e. in transverse configuration. In order to do so, the dependence on each parameter as predicted by the model, according to Eq. 5.14 in Section 5.3.1, is compared to the experimentally measured dependence.

Frequency dependence

The most prominent characteristic of the Hooge model is that the noise power spectral density scales inversely with the frequency. The model can be fine-tuned by introducing an additional parameter, γ, used as exponent for the frequency, i.e. $N_{PSD} \propto 1/f^{\gamma}$. This exponent may vary around unity to some extent to optimize the fit to the data. In the following, we will not consider this additional exponent in order to reduce complexity. The noise voltage spectral density then exhibits a $1/f^{0.5}$ dependence according to the Hooge relation (see Eq. 5.14). Panel c of Fig. 5.6 reveals nicely such a dependence for the extracted $N_{V,1/f}$ data. It was obtained on a hBN-encapsulated device for fixed gate voltage and corresponds to a vertical line cut through the colormaps in Fig. 5.5. The CNP was chosen since the 1/f noise amplitude is the largest and 1/f noise prevails over the widest frequency range. However, we also found that the $1/f^{0.5}$ line fits well to the low frequency noise data for other gate voltages (not shown), and also in case of all other devices we measured.

Alternatively, $N_{V,1/f}$ as a function of gate voltage, or density, can be compared for different frequencies. This corresponds to different horizontal line cuts through the colormaps in Fig. 5.5. Examples are shown in panel b of Fig. 5.7. By applying a relatively high bias current, we ensured that 1/f noise prevails over the entire density range and up to several tens of kHz. Only at 50 kHz, thermal noise eventually becomes the dominant noise source at higher densities. The inset of panel b then shows plots of $N_{V,1/f}$ multiplied by $f^{0.5}$ revealing that all curves almost perfectly overlap. Our findings unambiguously demonstrate that the low frequency noise in graphene Hall elements, i.e. in transverse configuration, can be described by the 1/f dependence in the Hooge model. This allows us to easily extrapolate the noise voltage (or the magnetic field resolution) measured at a certain frequency to any other frequency, if 1/f noise dominates. This procedure is used in Sections 5.7.1 and 5.7.3 when simulating graphene Hall elements and comparing different reported values, respectively. Finally, the 1/f frequency dependence is the best studied property of

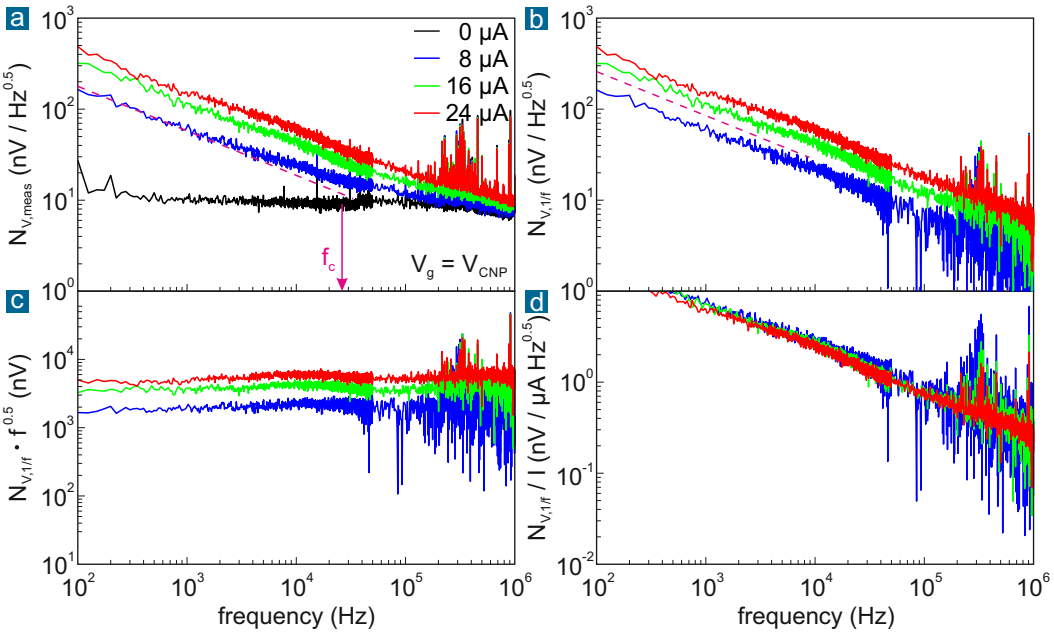

Figure 5.6. Frequency dependence of the noise voltage spectral density for different current values at fixed gate voltage (here, exemplarily, at the CNP where the 1/f noise amplitude has a maximum). Data was obtained on the same device as in Fig. 5.5, i.e. the shown curves correspond to a vertical line cut at $V_g = V_{CNP} = 0.5\,V$. a: Original measured data showing nicely the transition from a $1/f^{0.5}$-dependence to the white noise floor which allows to determine the corner frequency (dashed magenta line). b: The 1/f noise contribution obtained by using Eq. 5.24. c: Data from panel b multiplied by $f^{0.5}$ which demonstrates the dominance of 1/f noise over a large frequency range. d: Data from panel b divided by the current revealing a linear current dependence.

noise in graphene Hall elements and was also confirmed independently by several other research groups (Refs. 198, 348, 397, 398, 418).

Current dependence

The Hooge relation predicts a linear dependence of the noise voltage spectral density on the bias voltage. In our measurements, we operate our devices in constant current mode. For a given gate voltage (charge carrier density) and constant temperature, the two-terminal resistance value is a constant and, hence, the noise voltage is expected to scale linearly with the applied current (see Eq. 5.14). This is nicely confirmed by our $N_{V,1/f}$ data as a function of frequency for different current values in panel d of Fig. 5.6. Moreover, for a fixed frequency of 1 kHz, the measured noise voltage as a function of gate voltage for several current values is depicted in panel a

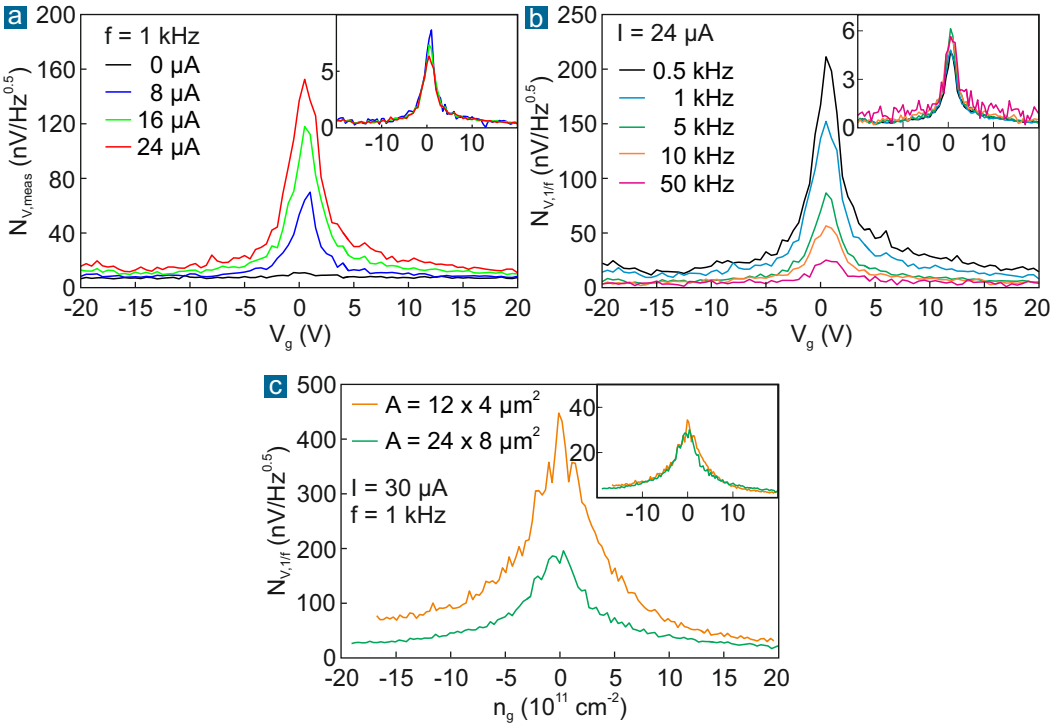

Figure 5.7. Dependence of the 1/f noise in graphene Hall devices on different parameters (in accordance to the Hooge model). a: Measured noise voltage as function of gate voltage for different current values at constant frequency (same hBN-encapsulated device as in Fig. 5.5). Inset: $N_{V,1/f}$ data normalized by the current, i.e. $N_{V,1/f}/I$ (in units of $nV/Hz^{0.5}\,\mu A$). The overlap of all curves confirms the linear dependence on the current. b: $N_{V,1/f}$ as function of gate voltage for different frequencies but constant current (same device as in panel a). Inset: Same data but multiplied by $f^{0.5}$, i.e. $N_{V,1/f} \cdot f^{0.5}$ (in units of μV), showing the characteristic frequency dependence of 1/f noise. c: $N_{V,1/f}$ versus charge carrier density for two graphene devices on SiO_2 with different dimensions but same geometry ($L/W = 3$). Inset: Normalized noise voltage as defined in Eq. 5.25 (in units of $nV\,\mu m/Hz^{0.5}\,\mu A$) confirming the areal dependence predicted by the Hooge model.

of Fig. 5.7. The 1/f noise component is extracted and plotted in the inset normalized by the current. Again, the different curves coincide almost perfectly[22]. The linear current dependence was also reported by other studies of graphene Hall elements

22 At the CNP, the data shown here may suggest that there is still a remaining current dependence. However, this is just a coincidence in this data set. For other frequencies, e.g. 3 kHz, or for other samples, the different curves either perfectly overlap at the CNP or show a different, random deviation of the curves at the CNP.

that measured the noise voltage in transverse configuration (Refs. 348, 397, 398, 418).

Dependence on area

The Hooge relation predicts that the bias voltage normalized 1/f noise voltage spectral density, $N_{V,1/f}/U$, scales inversely with the square root of the channel area, i.e. $\propto 1/\sqrt{A}$ where $A = L \cdot W$. When devices are operated in constant current mode, as in our case, the dependence of the measured noise voltage spectral density on the length L and width W of the channel becomes more complex. In addition, both the dependence of the channel resistance on the L/W ratio as well as the inverse dependence of the contact resistance on the width have to be taken into account (see Eq. 5.14 in Section 5.3.1). When the contact resistance contribution to the two terminal resistance is negligible, Eq. 5.15 in Section 5.3.1 applies. In this case, under the assumption that 1/f noise dominates and if the same frequency is considered in all measurements (e.g. 1 kHz), the noise voltage can be normalized with respect to the absolute dimensions, the geometry and the bias current as follows[23]:

$$N_{V,norm}(R_{ch} \gg R_c) = \frac{N_{V,1/f}}{I} \cdot \sqrt{A} \cdot \frac{W}{L} = \frac{N_{V,1/f}}{I} \cdot \sqrt{\frac{W}{L}} \cdot W\,. \tag{5.25}$$

In panel c of Fig. 5.7 we show the measured noise voltage spectral density as a function of density for two non-encapsulated graphene devices on SiO_2 with a notably different channel areas (factor of four) but same Hall cross geometry, i.e. L/W ratio. Both devices were fabricated from one single, very large graphene flake and underwent exactly the same device processing steps. The contact resistances as well as mobility values were almost identical. As discussed in more detail in the following subsection, non-encapsulated graphene devices on SiO_2 exhibit a larger 1/f noise level than hBN-encapsulated devices such that both devices were operated unambiguously in the 1/f noise regime over the entire density range. This was also

23 The same logic and Eq. 5.25 will be used in Section 5.7.3 to normalize different magnetic field resolution values (Eq. 5.28), in order to compare our results with the findings of other studies of graphene Hall elements in the literature. In the 1/f noise regime, B_{min} is independent from the current. Thus, a normalization by multiplying with $\sqrt{W/L} \cdot W$ is sufficient. Moreover, if the noise voltage spectral density values were obtained at different frequencies, one needs to extrapolate these values to a common reference frequency, e.g. 1 kHz, using the Hooge relation. We again emphasize that pre-requisites for performing such normalization and comparison are the following: i) each device is operated in the 1/f regime, ii) the contact resistances are negligible and iii) the best obtainable B_{min} value over the accessible density range is reported for each study (see Section 5.7.3).

verified by studying the current and frequency dependence analogous to the hBN-encapsulated device in panel a and b of the same figure (data not shown). Moreover, the contact resistance contribution is less relevant since the channel resistance is higher due to the lower mobility. Hence, Eqs. 5.15 and 5.25 above are applicable. The result is shown in the inset of panel c in Fig. 5.7. Even though operated in constant current mode, both curves overlap nicely after normalization by $\sqrt{\mathrm{A}}$ in accordance with the Hooge relation. This is also supported by works from another research group where the authors attribute the reduction of the noise level in their graphene Hall elements to the notable increase in the device area [348, 399].

Alternatively, the noise voltage spectral density can be divided by the bias voltage which we recorded in parallel in all our measurements. Note that the bias voltage is a function of density for constant current mode. In doing so, the complexity introduced by the two terminal resistance in Eq. 5.14 is avoided. This allows verifying the $\propto 1/\sqrt{\mathrm{A}}$ dependence more straightforwardly. This analysis is indirectly conducted in the following Section 5.5.3 for a large set of devices when extracting the Hooge parameter using Eq. 5.26. In the single carrier regime, away from the CNP, the extracted α_{H} values for all devices on SiO_2 are around 10^{-3} despite the large variation in channel area between 40 and 1600 $\mu\mathrm{m}^2$. This finding again supports the Hooge relation. Finally, we emphasize that graphene devices on SiO_2 are quite comparable in terms of mobility and the contact resistance contribution is in general less relevant (except for devices with very small area). However, in case of hBN-encapsulated devices, mobility and contact resistance values varied substantially among different samples. However, a substantial difference in device quality, i.e. mobility, has a notable impact on the α_{H} value and a larger contact resistance contribution increases the inaccuracy when determining the Hooge parameter (see detailed discussion in Section 5.5.3). Moreover, the maximum area of hBN-encapsulated devices is limited due to the need for stacking several exfoliated flakes on top of each other. This does not allow varying the area over a wide range. Hence, the data on hBN-encapsulated devices cannot be considered for a proper study of the dependence on area.

Density dependence

As discussed in Section 5.3.1, the density dependence in case of constant bias current operation is more complex than the $1/\sqrt{\mathrm{n}}$ dependence which follows directly from the Hooge model (see Eq. 5.14 in Section 5.3.1). This is because of the additional density dependence of the two-terminal resistance. The bias voltage becomes a function of density for constant current operation. The two terminal resistance is composed of the channel resistance and the contact resistance. The channel resistance exhibits both a direct 1/n dependence and an indirect density dependence through $\mu(\mathrm{n})$. The contact resistance has a sample-specific density dependence, as

discussed in the previous sections. Moreover, the Hooge model covers strictly speaking only the single charge carrier type regime[24].

In order to reduce the aforementioned complexity, we suggest to analyze the bias voltage normalized noise voltage spectral density, i.e. $N_{V,meas}(n)/U(n)$, since the bias voltage was recorded in all our measurements. As will be shown and discussed in detail in the following subsection 5.5.3, the Hooge parameter α_H can be extracted from the measured noise voltage spectral density using Eq. 5.26. The results are shown in Figure 5.8 and reveal that the obtained Hooge parameter is almost independent of the density (in the single carrier type regime, away from the CNP)[25]. If at all, the Hooge parameter slightly increases at higher density. This can be explained by the lower mobility at higher density while the Hooge parameter is known to be larger for lower mobility, as discussed in Section 5.3.3. To conclude, the finding that the extracted Hooge parameter does not or only weakly depend on density supports the Hooge model and its prediction about the density dependence.

It should be pointed out that, for all studied devices, we have observed a Λ-shaped density dependence of both the measured as well as the bias voltage normalized noise voltage. This means that the noise voltage has a maximum at the CNP and then steadily decreases towards higher densities in accordance with the Hooge relation. Our finding is in agreement with other studies conducting noise voltage measurements in the transverse configuration (Refs. 395, 397). A V-shaped or M-shaped dependence, which sometimes was reported by other studies in case of longitudinal measurements, was never observed. As discussed in Section 5.3.3, this suggests that resistance fluctuations prevail.

Dependence on device quality (charge carrier mobility and contact resistivity)

Finally, in order to provide a complete picture, the Hooge parameter α_H in the model (Eq. 5.14) was found to depend on the device quality, i.e. the mobility (for instance, Ref. [404] and references therein). For devices with substantially larger mobility, the Hooge parameter and, accordingly, the 1/f noise amplitude is lower. This will be demonstrated for a large ensemble of devices in the following subsection. Moreover, as directly follows from Eq. 5.14, the contact resistivity contributes

24 Note that we already modified the Hooge model in Eq. 5.14 by taking the total charge carrier density into account. In case of graphene, the total charge carrier density can be described by the channel approximation (5.12). This will be used in our simulation of graphene Hall elements in Section 5.7.

25 Note that Eq. 5.26 does not consider the coexistence of electrons and holes at the CNP and is only applicable in the single carrier regime. Hence, the Hooge parameter values around the CNP must be neglected.

potentially stronger to the two-terminal resistance and, accordingly, to the 1/f noise in case of devices with high mobility than in case of devices with low mobility. This will be also discussed further in the following sections.

5.5.3. Substrate and layer dependence of the Hooge parameter

As demonstrated above in the previous subsection, the noise voltage between the Hall contacts in graphene Hall transducers, i.e. the noise voltage in transverse configuration, can be described at low frequencies by the empirical Hooge model for 1/f noise. The measured noise voltage spectral density at lower frequencies, typically below 10 kHz, obeys the characteristic frequency dependence, scales linearly with the applied current and depends inversely on the square root of the channel area. As discussed in Section 5.3.2, the charge carrier density dependence of the measured noise voltage for constant bias current is more complex than the direct $1/\sqrt{n}$ dependence in the Hooge formula. This is due to the density dependence of the two terminal resistance between the source and drain contacts, R(n) in Equation 5.11, which is composed of channel and contact resistances. Hence, the measured noise voltage must be normalized by the density dependent bias voltage in case of measurements with constant current, $U(n) = R(n) \cdot I$, in order to eliminate these indirect density dependencies. In our measurements, which were all performed with a constant bias current, we always recorded the two terminal voltage between the source and drain contacts while tuning the density. The Hooge parameter can then be obtained from the normalized noise voltage spectral density at low frequency, i.e. in the 1/f noise regime, using the following re-arrangement of the Hooge formula in Eq. 5.14 and considering only the single carrier regime with the gate controlled density n_g:

$$\alpha_H = \left(\frac{N_{V,1/f}}{U}\right)^2 f\, A\, n_g \ . \tag{5.26}$$

For the statistics based on a large set of devices, presented here, we used noise voltage spectral density data at $f = 1\,\mathrm{kHz}$ and applied a high enough current such that the noise voltage was notably larger than the thermal noise floor over the entire density range, i.e. larger than the noise voltage without applying a current. We then obtained $N_{V,1/f}$ by subtracting the noise floor from the measured noise voltage data for the respective current value (see Eq. 5.6 and Section 5.3.1). Since the absolute dimensions and geometries varied among the investigated devices, we

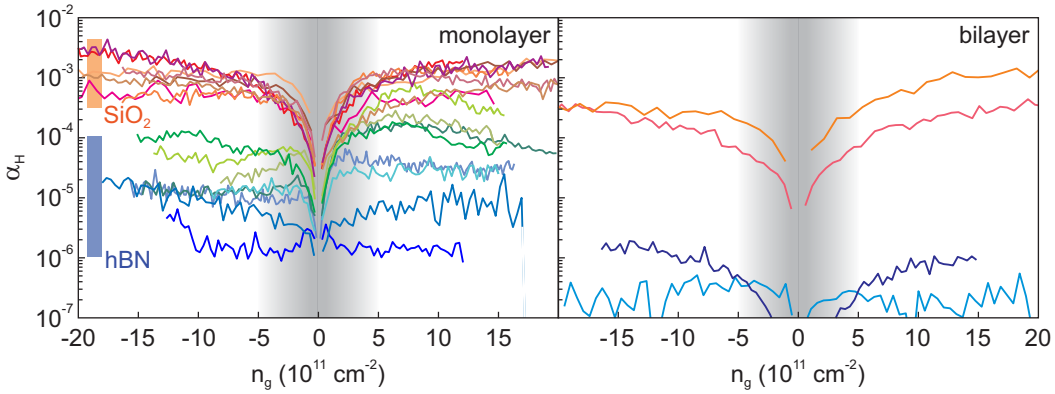

Figure 5.8. From noise voltage measurements in transverse configuration extracted Hooge parameters as a function of the gate controlled carrier density (using Eq. 5.26). Each line corresponds to an individual sample. We distinguish monolayer (left) and bilayer devices (right panel) as well as devices on SiO_2 (different orange, red or violet line colors) and hBN-encapsulated devices (different green or blue line colors). See main text for further discussion.

chose the current in each case such that the maximum bias voltage values were comparable in magnitude[26]. Typically, bias currents between 20 and 40 μA were used. Each device's channel area was known from design and was double-checked by optical or AFM images. For hBN devices, the channel areas ranged from 7 to 48 μm^2, while devices on SiO_2 ranged from 40 to 1600 μm^2 (also two CVD grown graphene devices on SiO_2 were studied which allowed the fabrication of devices with larger channel area).

Substrate dependence

The left panel of Figure 5.8 shows the obtained α_H values versus the gate controlled carrier density, n_g, for a multitude of monolayer graphene devices. We distinguish between bare graphene devices on SiO_2 (different orange, red or violet line colors) and hBN-encapsulated devices (different green or blue line colors). The greenish colored data was recorded on heterostructures that were not yet annealed at high temperatures after stacking and/or not treated by the AFM technique. Some of the

26 That means a bias voltage of at most a few hundred mV at the CNP. However, this constitutes only a precautionary measure to avoid other effects such as a too large potential gradient along the device channel. We analyzed the obtained data for each sample for various frequencies, e.g. 500 Hz or 2 kHz, and different currents, i.e. bias voltages, and found no notable difference in $\alpha_H(n_g)$ - as it also should be the case for 1/f noise according to the Hooge model.

latter devices exhibit an asymmetry between the electron and hole side which is introduced when dividing through the two-terminal bias voltage (the original N_V data is more symmetric). Data on hBN-encapsulated graphene devices fabricated following our optimization procedure, which we described in Section 3.1.5, is plotted using bluish line colors. Moreover, the above formula (Eq. 5.26) strictly only holds in the high density regime with only one charge carrier type but not in the vicinity of the CNP where electrons and holes coexist (gray shaded area must be disregarded)[27]. Considering the fluctuations in the data, the Hooge parameter in this single carrier regime away from the CNP does not or only weakly depend on the density. This finding is in agreement with the empirical model which assumes the Hooge parameter to be constant. The Hooge parameter, however, is known to depend on the material's quality, of which the mobility is a proper indicator (see Ref. 404 and references therein). In graphene, the mobility does not change substantially with density over a wide density range (see Fig. 4.3 in Section 4.1.4). Hence, each data set showing α_H to be almost independent of the density supports the applicability of the empirical Hooge model. From Equation 5.26 then follows that $N_{V,1/f}/U \propto 1/\sqrt{n_g}$ if one would operate the element in constant voltage mode, while the measured noise voltage for constant current exhibits a more complex density dependence, as discussed above.

Despite the large differences in channel area, the $\alpha_H(n_g)$ curves are very similar for graphene on SiO_2 devices. This again proves the Hooge model's validity in terms of the channel area dependence. We find $\alpha_{H,SiO_2} \approx 10^{-3}$ away from the CNP (ranging from $5 \cdot 10^{-4}$ to $3 \cdot 10^{-3}$). This is in good agreement with elsewhere reported Hooge parameters for graphene on SiO_2, obtained either from current or voltage fluctuation measurements along the channel, i.e. in longitudinal configuration [392, 403, 404, 411, 419], or from noise voltage measurements between the Hall contacts, i.e. in transverse configuration [398, 399]. $\alpha_H \approx 10^{-3}$ also constitutes a typical value for other conventional, bulk materials (Refs. 404, 420–422).

Zhang and coworkers (Ref. 404) studied voltage fluctuations along small but free-standing two-terminal graphene devices, possessing a much higher mobility than SiO_2 supported devices. They found lower Hooge parameters between 10^{-4} and 10^{-5}. Moreover, their data suggests a slope between the Hooge parameter and the mobility of -3 for SiO_2 supported and -1.5 for suspended devices. That means for the latter case that an increase in mobility by one order of magnitude lowers

27 In analogy to the Hall constant modeling in Section 4.1.2, one could replace n_g by $\sqrt{n_{t,0}^2 + n_g^2}$ to account approximately for the two carrier regime.

the Hooge parameter by 1.5 orders of magnitude[28]. For our hBN-encapsulated devices, we obtain Hooge parameters in the single carrier regime between 10^{-4} and even 10^{-6}. We find the lowest Hooge parameters for heterostructures that were fabricated using our proposed procedure involving both high temperature annealing after stacking and AFM ironing (see Section 3.1.5). As demonstrated in detail in the previous chapters, these devices exhibit the highest quality and mobility (measured in four terminal configuration).

However, a quantitative investigation of the mobility dependence of the Hooge parameter remains challenging (and should be addressed in future studies). Several of the hBN-encapsulated devices, studied here, were fabricated in Hall cross geometry, allowing only for a two-terminal resistance measurement. Since contact resistances are not negligible in case of our few micrometer sized devices and vary among the devices, the extracted mobility values using the Drude model are inaccurate. This impedes a proper quantitative statement about $\alpha_{\mathrm{H}}(\mu)$ when comparing our different samples that were designed primarily for Hall sensor performance studies. Moreover, the equations to calculate the Drude mobility from the resistivity and the above Hooge parameter formula are only valid in the regime with a single carrier type. At the crossover to the two carrier regime, this leads to a steep increase of the Drude mobility and decrease of the Hooge parameter, respectively. As shown in Figure 4.3 in Section 4.1.4, other mobility definitions (two carrier Drude model and Hall mobility) can be used. When doing so, the density dependence of the mobility is comparatively small over the entire density range. For the hBN-encapsulated device presented in Figure 4.3, the mobility increases from about $50,000\,\mathrm{cm}^2/\mathrm{Vs}$ at high density to around $70,000\,\mathrm{cm}^2/\mathrm{Vs}$ at the CNP (using the two carrier Drude definition). This does not correspond to a change over at least one order of magnitude which appears to be necessary in order to obtain a clear decrease in α_{H} considering the fluctuations in the data. Therefore, we restrict us to a more qualitative discussion of α_{H} away from the CNP, i.e. for $\mathrm{n_g} > 10^{11}\,\mathrm{cm}^{-2}$.

Our data presented in Figure 5.8 suggests that the Hooge parameter for hBN-encapsulated devices is on average around two orders of magnitude smaller than for

28 We note that in the work of Zhang et al. (Ref. 404) the Hooge parameters varied notably between different devices, like in our case. The slope value was determined by tuning the density of each device in order to vary also the mobility through exploiting the relatively weak $\mu(\mathrm{n})$ dependence. The mobility, however, was calculated from the measured resistivity. This leads to a $\alpha_{\mathrm{H}}(\mu(\mathrm{n_g}))$ behavior comparable to our data in Fig. 5.8. As already discussed, Eq. 5.26 strictly holds only in the regime with a single carrier type. Moreover, contact resistance distributions are relatively large in small devices with two-terminal geometry. This also leads to inaccuracies.

SiO_2 supported devices[29]. While $\alpha_{H,SiO_2} \approx 10^{-3}$, $\alpha_{H,hBN} \approx 10^{-5}$ seems to be a conservative value that can be reached reliably using our post-processing improvement steps (bluish line colors). A decline in the Hooge parameter by two orders of magnitude corresponds to a decrease of the measured noise voltage spectral density at low frequencies by about one order of magnitude (since $N_{V,1/f} \propto \sqrt{\alpha_H}$). This agrees well with findings from other groups which studied current or voltage fluctuations in longitudinal configuration and compared graphene on SiO_2 and hBN substrates (see Refs. 406, 407 and Section 5.3.3). However, the Hooge parameter was not determined in these studies. In their recent work, Kakkar and coworkers (Ref. 409) investigated hBN-encapsulated graphene FETs and found the Hooge parameter to be around $5 \cdot 10^{-6}$ when using graphitic gates. Behera *et al.* (Ref. 417) reported a Hooge parameter of 10^{-5} for hBN-encapsulated graphene which was obtained from longitudinal resistance noise measurements. Unfortunately, recent studies on hBN-encapsulated graphene Hall transducers do not provide a detailed analysis of the measured noise voltage between the Hall contacts (if detailed N_V data is shown at all). Neither do they compare between hBN and SiO_2 substrates regarding $N_{V,1/f}$ or α_H values [198, 349, 397]. The latter publications, however, provide B_{min} values for hBN-encapsulated devices which are comparable to our results, as discussed in the next section. Hence, it is reasonable to assume that extracting the Hooge parameter from their original noise voltage data would result in similar values as in our measurements. Moreover, our results are consistent with studies on other material systems. It was as well found that the Hooge parameter can be two or three orders of magnitude smaller for high quality materials than the typical value of $\approx 10^{-3}$ [420, 422–424]. For example, noise voltage measurements both in longitudinal and transverse configuration on MBE grown InSb devices result in Hooge parameter values around $2 \cdot 10^{-5}$ [423, 424].

The above detailed analysis of our noise voltage measurements in transverse configuration, which was performed on a multitude of different graphene devices, demonstrates nicely the applicability of the empirical Hooge model. Through encapsulation with hBN the noise level in our few micrometer sized graphene devices can be reduced by about two orders of magnitude compared to devices on SiO_2 that are not

29 This is also plausible from the following back of the envelope calculation. Room temperature mobility values are typically in the range of $3,000 - 5,000\,cm^2/Vs$ on SiO_2, while for high quality hBN-encapsulated devices around $60,000 - 100,000\,cm^2/Vs$ are reached. This corresponds to a factor of 20. Assuming a factor of -1.5 between Hooge parameter and mobility, as reported by Zhang et al. for high quality graphene (Ref. 404), this results in a reduction of $\alpha_{H,SiO_2} \approx 10^{-3}$ by about 30 which results in $\alpha_{H,hBN} \approx 3 \cdot 10^{-5}$.

encapsulated ($\alpha_{H,hBN} \approx 10^{-5}$). Knowing the Hooge parameter in transverse configuration allows us to predict properly the noise level for arbitrary operating conditions, i.e. for typical current and operating frequency values in commercial applications. Moreover, we can estimate the performance of large-area, hBN-encapsulated graphene Hall transducers having a footprint similar to today's commercial Hall elements, i.e. on the order of $100\,\mathrm{x}\,100\,\mu\mathrm{m}^2$ (see Section 5.7.2). Of course, this requires solving the remaining challenges in the fabrication of such large-area hBN-encapsulated devices with similar quality as exfoliated flakes (see Section 5.8). Moreover, noise arises from both channel and metal-graphene contact areas [409]. Recent studies found that 1/f noise in exfoliated, few micrometer sized devices is governed to a notable extent by resistance fluctuations at the contacts [412, 413, 425]. Our hBN-encapsulated devices were not optimized in this regard. Actually, we estimate the contact resistivity, ρ_c, in the 10^{11} to $10^{12}\mathrm{cm}^{-2}$ density range to be between several hundred and up to $3000\,\Omega\,\mu$m when comparing four terminal versus two terminal resistance measurements in Hallbar devices. Wang and coworkers (Ref. 10) report a contact resistivity for this range between 200 and $1000\,\Omega\,\mu$m in their publication that originally introduced the edge-contact geometry for hBN-encapsulated graphene devices. Obviously, the contact resistances may vary substantially depending on the used etching and evaporation systems and recipes. We used similar recipes but a different dry etching system. This leaves room for further improvement of graphene Hall transducers in commercial production, i.e. by optimizing contact resistances. Nevertheless, we believe that $\alpha_{H,hBN} \approx 10^{-5}$ is a realistic value to calculate with and a feasible target for large-scale fabrication.

Layer dependence

To our knowledge, bilayer or multilayer graphene Hall transducers have not been studied so far in the literature. Accordingly, no data on noise voltage spectral density in transverse configuration is available. However, 1/f noise in bilayer devices was measured in longitudinal configuration both for suspended and SiO_2 supported devices [403, 410–412]. Liu and coworkers (Ref. 410) even investigated the atomic layer dependence of the conductance noise from single to > 10 atomic planes and found that 1/f noise is a surface phenomenon below seven layers (above seven layers volume noise begins to dominate), as already discussed in Section 5.3.3. A $1/N^2$ scaling could be extracted from their data, i.e. conductance noise in bilayer (N=2) was about one fourth of monolayer graphene. We measured Hall voltage noise, i.e. in transverse direction, for four bilayer devices, two on SiO_2 and two encapsulated between hBN. We verified the validitiy of the Hooge model as in the previous section for monolayer, i.e. $N_{V,1/f}$ scales with the applied current and shows characteristic

frequency dependence (not shown here). The Hooge parameters obtained from Eq. 5.26 are provided in the right panel of Figure 5.8. The devices on SiO_2 do not differ notably from the best monolayer devices. In case of hBN-encapsulation, however, we find the Hooge parameter for bilayer graphene to be at least one order of magnitude smaller than in case of the best hBN-encapsulated monolayer devices ($\alpha_{\mathrm{H,hBN}}(\mathrm{bi}) \approx 10^{-6}$).

Since $N_{V,1/f} \propto \sqrt{\alpha_H}$, the noise voltage spectral density at low frequency is reduced by at least a factor of three when switching from monolayer to bilayer graphene Hall devices for otherwise same geometries and operating parameters. This means that the numerator in the equation for the magnetic field resolution in the 1/f noise regime is reduced by this factor, accordingly (Eq. 5.7 in Section 5.2.2). However, also the maximum Hall coefficient, i.e. the denominator in the equation, is smaller for bilayer devices compared to monolayer devices, as we have shown in Figure 4.2 in Section 4.1.3. For hBN-encapsulated bilayer devices, we found $R_{H,max} \approx 1000\,\Omega/\mathrm{T}$ at room temperature, while hBN-encapsulated monolayer devices exhibit reliably five times larger maximum Hall coefficients. Hence, the minimum detectable magnetic field in case of an operation in the 1/f noise regime is in total larger for bilayer devices, i.e. the magnetic field resolution is worse (by a factor of 5/3 using the above factors). As discussed in the following section in more detail, the best magnetic field resolution is not obtained exactly at the density for which the maximum Hall coefficient is obtained, but for slightly higher densities. Our argumentation, however, remains valid since Hall coefficient values are also at this ideal working point still larger for monolayer devices. Only at high densities the $R_H(n_g)$ curves for monolayer and bilayer graphene eventually converge and the lower 1/f noise in bilayer devices is reflected in a better magnetic field resolution for operation in the 1/f noise regime. An advantage, however, is that the smaller 1/f noise in bilayer devices results in smaller corner frequencies f_c. This allows for an operation in the thermal noise regime at lower frequencies.

Finally, the above argumentation comparing monolayer and bilayer devices also holds for multilayer devices with three or more layers. Even though the noise level drops further with increasing layer thickness, the reduction in sensitivity with layer thickness is larger than the improvement of the noise level. As a result, the obtainable magnetic field resolution is worse than for monolayer graphene devices. To conclude, bilayer or multilayer graphene Hall transducers do not outperform monolayer devices and are more challenging to fabricate in large scale (e.g. by CVD growth). Therefore, we focus on monolayer devices in the following.

5.6. Magnetic field resolution of graphene Hall transducers

According to our measurement scheme in Section 5.4, for each gate voltage step the current-related sensitivity, i.e. the Hall coefficient, is determined from the measured $U_H(B)$ curves and then the noise voltage spectrum is recorded for several different currents. Combining these two values allows us to determine the magnetic field resolution as follows (see also Eq. 5.16 in Section 5.3.2):

$$B_{min,(meas)}(I) = \frac{N_{V,meas}(I)}{|R_H \cdot I|} \ . \tag{5.27}$$

An example of this procedure is given in the left panels (a,c,e) of Fig. 5.10. The obtained magnetic field resolution as a function of both gate voltage and frequency is presented in the colormaps of Fig. 5.9 for two different current values. The data reveal that the best magnetic field resolution, i.e. the smallest B_{min} value, is obtained in the intermediate density range in case of the hBN-encapsulated graphene device with few square micrometer area studied here. For higher density, the B_{min} value increases again. Moreover, increasing the bias current results in smaller B_{min} values only in the intermediate to high frequency regime while at lower frequencies B_{min} already saturates for low current values.

In the first subsection, 5.6.1, all inter-dependencies between the different operating parameters and the interplay of the two dominating noise mechanisms, 1/f and thermal noise, will be studied. Finally, in the second subsection, 5.6.2, the temperature dependence of the magnetic field resolution will be discussed in detail. It is a combination of the temperature dependence of the Hall coefficient and the noise voltage, as well as the thermal drift of the charge neutrality point.

5.6.1. Interplay of operating parameters and dominating noise mechanisms

The dependence of the noise voltage on the operating parameters has been studied thoroughly in the preceding Section 5.5. For a given gate voltage, i.e. charge carrier density working point, the crossover from 1/f noise in the low frequency regime to thermal noise in the high frequency regime is determined by the bias current. In

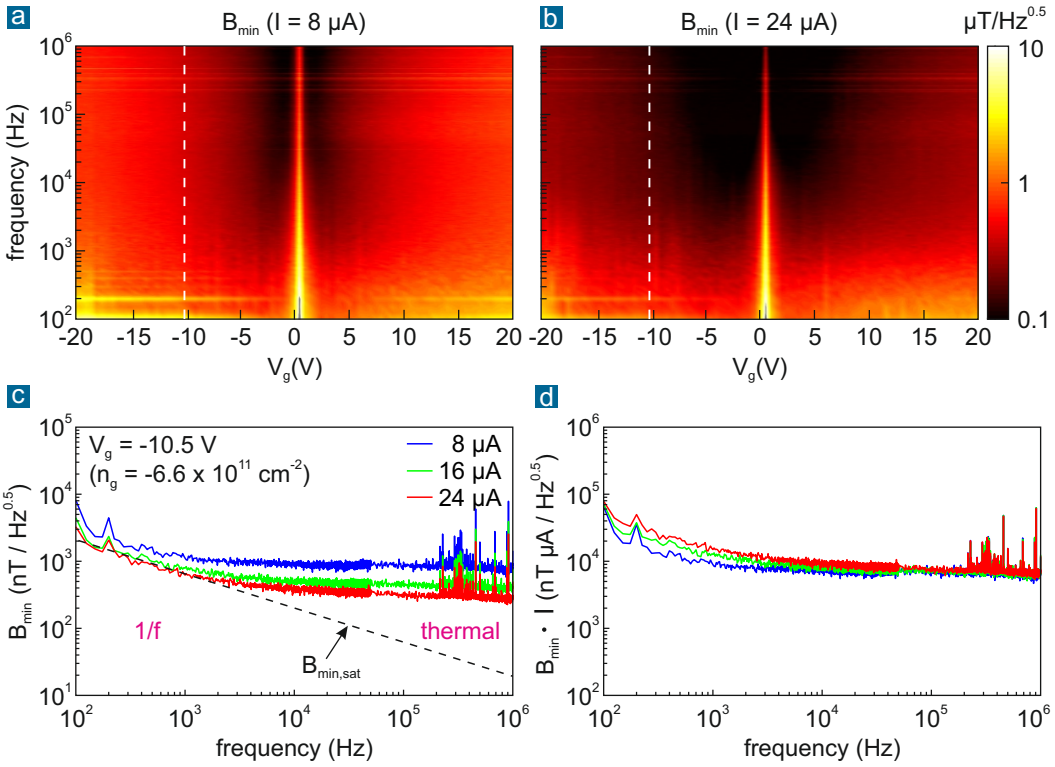

Figure 5.9. Magnetic field resolution as function of gate voltage and frequency (for the same hBN-encapsulated graphene device as in Fig. 5.5 in Section 5.5). a,b: B_{min} for two different current values, obtained from the measured noise voltage and the current-related sensitivity, i.e. the Hall coefficient, using Eq. 5.27. c: B_{min} versus frequency for fixed gate voltage and three different current values. For $8\,\mu$A and $24\,\mu$A, these data correspond to the white dashed lines in panel a and b, respectively. In the low frequency regime, 1/f noise dominates and the magnetic field resolution saturates, i.e. increasing the current no longer improves B_{min}. In the high frequency regime, the thermal noise floor is reached and B_{min} scales inversely with the applied current. d: Same data as in panel c but multiplied by the applied current in order to demonstrate the 1/I dependence in the thermal noise regime.

case of the magnetic field resolution, however, an additional current dependence is introduced by the denominator in Eq. 5.27. Moreover, the Hall coefficient depends on the density as discussed in Chapter 4. This leads to a different interplay of the operating parameters for the magnetic field resolution. In Section 5.3.2 , two different regimes were already discussed theoretically. They will be confirmed by the following measurements.

Thermal noise regime

When the device is operated in the thermal noise regime, the noise voltage (the nominator in Eq. 5.27) neither depends on the current, nor on the frequency. Hence, B_{min} scales inversely with the current. This is apparent in Fig. 5.9 (in particular in panel d). Moreover, panels d and f of Fig. 5.10 show that as soon as the thermal noise regime is reached, corresponding to the gate voltage range shaded in turquoise, increasing the frequency no longer improves the magnetic field resolution. When thermal noise is the dominant noise source, one would expect an increase of the B_{min} value with density according to Eq. 5.17 and, in particular, Eq. 5.18 in Section 5.3.2 which describes the case when the contact resistance contribution is negligible. This means that operating the device at low density is preferable when only considering thermal noise (and not taking into account the coexistence of two charge carrier types at the CNP which eventually reduces the Hall coefficient). However, in real devices without applying particular 1/f noise reduction techniques, such as the spinning current method, the low frequency and low density range is typically governed by 1/f noise.

1/f noise regime

For any given density and frequency, the above approach to simply increase the current in order to improve the magnetic field resolution only works until 1/f noise becomes the prevailing noise source. Then, B_{min} saturates since both the nominator and the denominator in Eq. 5.27 scale linearly with the current. $B_{min,1/f} = B_{min,sat}$ no longer depends on the current (see also Eq. 5.19 in Section 5.3.2) and a further increase in current no longer results in an improvement of the magnetic field resolution. This means that for the given combination of density and frequency this $B_{min,sat}$ value is the best obtainable magnetic field resolution. In case of our few micrometer sized, hBN-encapsulated devices, 1/f noise typically dominates over the entire studied density range for frequencies below a few 10 kHz and even when applying only smaller current values. This is seen in panel e of Fig. 5.10 where the B_{min} values no longer change when moving from 16 μA to 24 μA current. In contrast, at a higher frequency of 50 kHz, B_{min} does not yet saturate at higher density values (panel b). This is also supported by the plots in panel d and f. In the magenta shaded gate voltage range and for frequencies at and below 10 kHz, the magnetic field resolution does not improve notably when increasing the current while a substantial improvement is found at higher density (in the thermal noise regime, as discussed above).

Best obtainable magnetic field resolution within the operating parameter space

From the previous paragraph follows that no general best magnetic field resolution value can be stated for a particular density working point alone. Instead, the obtainable B_{min} value will depend on potential restrictions of the operating frequency and/or bias current and must be always stated together with these parameters. As shown in panel c of Fig. 5.10 and also in Refs. 198, 348, 399, $B_{min,1/f} = B_{min,sat}$ scales with $1/f^{0.5}$ (black dashed line). However, an ever larger current must be applied in order to reach this optimal value for the given operating frequency. This is also demonstrated nicely in Ref. 398. In real applications, however, restrictions in terms of maximum current or power consumption as well as operating frequency will apply. A too large bias current will heat up the device, eventually destroying it. Moreover, a too large longitudinal voltage drop will also lead to doping asymmetries and a reduction of the Hall coefficient. A higher operating frequency requires appropriate performance of the other electronic components, increasing complexity and cost. This will be taken into account in our detailed discussion about the best obtainable performance of graphene Hall elements in the following Section 5.7.

Hence, it is possible to determine and state the best obtainable magnetic field resolution for a certain, fixed density working point when either restricting the maximum frequency and/or current. This is for instance the case for graphene devices on SiC with fixed doping level (Refs. 398, 418). However, the analysis becomes much more complex when also the density working point can be chosen arbitrarily, i.e. the question is what is the global best magnetic field resolution. For our few micrometer sized, hBN-encapsulated graphene Hall elements, we found that the lowest B_{min} values are obtained at intermediate density values, away from CNP, before B_{min} eventually increases again at higher densities (see colormaps in Fig. 5.9 and panels b and d-f in Fig. 5.10). As also demonstrated in Fig. 5.10, this optimal density range, n_{opt}, does not coincide with the density working point for which the Hall coefficient reaches its maximum. $R_{H,max}$ is obtained when the gate controlled density equals the residual total charge carrier density, i.e. $n_g = n_{t,0}$, as discussed in the preceding chapter. Instead, we find n_{opt} in the range of $2.5 - 5 \cdot 10^{11}\,cm^{-2}$ for all studied hBN-encapsulated graphene devices, which equals about 4 to $5 \cdot n_{t,0}$. Moreover, as will be discussed in the following Sections 5.6.2 and 5.7 in more detail, this intermediate density regime is also favorable to reduce the impact of temperature drift and other Hall transducer performance indicators such as non-linearity and offset (see also Figs. 5.11 and 5.12). Unfortunately, not many other studies in the literature investigated and discussed the density dependence of the noise voltage and the magnetic field resolution in detail. The data provided by Refs. 198 and 397 (both for hBN-encapsulated devices), however, also suggests that the density working point

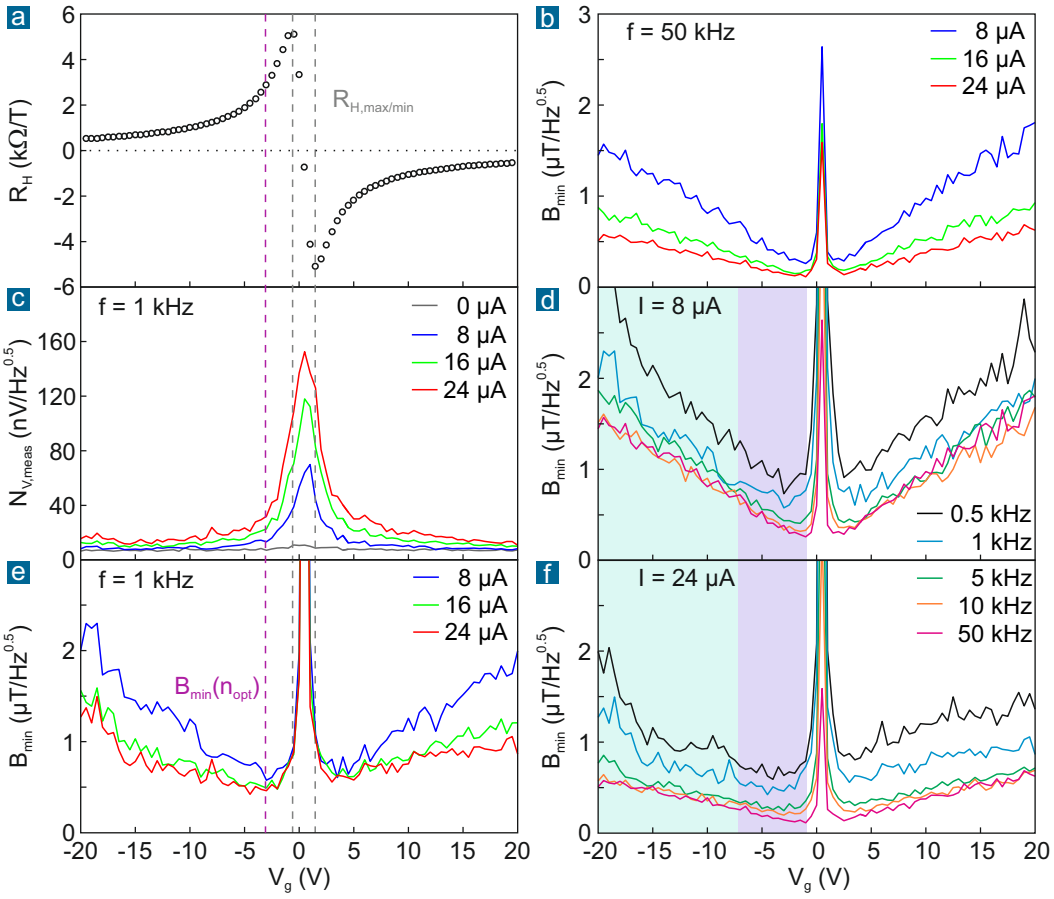

Figure 5.10. Magnetic field resolution as a function of gate voltage for different combinations of current and frequency (for the same hBN-encapsulated device as in Fig. 5.5). The measured Hall coefficient (a) and noise voltage spectral density (c) are used to determine the magnetic field resolution (e) using Eq. 5.16. At a low frequency of 1 kHz (e), 1/f noise dominates already over the entire density range for small current values and $B_{min,sat}$ is reached quickly. In contrast, at higher frequency (50 kHz), B_{min} does not yet saturate at higher density within the applied current range (b). Finally, panels (d) and (f) provide further insight into the interplay between 1/f noise (shaded in purple) and thermal noise (shaded in turquoise), as discussed in the main text.

for the best magnetic field resolution does not coincide with that of the highest sensitivity (largest Hall coefficient).

In case thermal noise prevails, a lower density working point is favorable to improve the magnetic field resolution. As follows from Eqs. 5.17 and 5.18 in Section 5.3.2,

B_{min} increases with density (assuming that the mobility is almost independent of density). For a negligible contact resistance contribution, one expects a $\sqrt{n}$ dependence. For the more realistic case that 1/f noise prevails, as is the case for our few micrometer sized devices, the density dependence of $B_{min,1/f} = B_{min,sat}$ is described by Eqs. 5.19 and 5.20 in Section 5.3.2 (considering the single carrier regime and constant current operation). $B_{min,1/f}$ decreases with increasing density when the contact resistance contribution to the two terminal resistance is negligible. If the latter is not negligible, there is a second part in Eq. 5.17 that scales with $\rho_c \cdot \sqrt{n}$, where ρ_c is the contact resistivity. This means that $B_{min,1/f}$ may also increase with density, depending on the contact resistivity value.

To conclude, the applied currents in our measurements of few micrometer sized devices are large enough such that 1/f noise dominates over the entire gate voltage or density range at lower frequencies (below 10 kHz). As discussed above, we find B_{min} to increase with density after reaching a minimum at intermediate density. Both the transition to dominating thermal noise at higher densities as well as the non-negligible contact resistance result in such a characteristic density dependence for our micron sized devices with high mobility. This is also demonstrated nicely in our simulations of graphene Hall elements which will be presented in the following Sections 5.7.1 and 5.7.2. As shown in Fig. 5.13 (panel c), we find a qualitatively similar behavior in our simulation results for larger contact resistivity values, i.e. B_{min} increases with density after reaching a minimum at intermediate density. As a consequence, the contact resistivity constitutes one of the four design parameters which we identified to be crucial to properly simulate and predict the performance of graphene Hall elements (see Table 5.2 in Section 5.7.1).

When the contact resistance contribution is negligible, e.g. for large-area devices and/or for devices with lower mobility such as on SiO_2 or SiC, the magnetic field resolution no longer increases with density. Instead, it continuously decreases with increasing density, however rather slowly, as demonstrated in our simulations of large-area devices in Fig. 5.14 in Section 5.7.2. Accordingly, operation at high density is favorable (but requires an ever larger current such that $B_{min,1/f} = B_{min,sat}$ is reached). The density dependence of B_{min} in our measurements on SiO_2 supported devices with larger area already showed the onset of such reversal of the density dependence (data not shown). This potentially also explains why large-area devices on SiC achieve promising absolute B_{min} values (see Refs.398, 418), even though they are operated at extremely high density for which the Hall coefficient is very small. However, after normalization of these absolute values, the obtained $B_{min \cdot W}$ values are still notably higher, i.e. worse, than for high quality, hBN-encapsulated devices (see Table 5.3 and the detailed discussion in Section 5.7.3). For our few micrometer sized, hBN-encapsulated graphene Hall elements, the best obtainable magnetic field

resolution, i.e. $B_{min,1/f} = B_{min,sat}$ at n_{opt}, is about $200 - 500\,nT/\sqrt{Hz}$, depending on the exact device dimensions. We note again that only a bias current of few tens of μA must be applied in order to reach these values. For $B_{min \cdot W}$ we obtain approximately $500 - 1500\,nT\,\mu m/\sqrt{Hz}$ which is comparable to the results in Ref. 397 (which even report slightly smaller values as summarized in Section 5.7.3).

5.6.2. Temperature dependence

As described in Section 5.4, our measurement setup is equipped with a ceramic heater and reference temperature sensor. This allows us to study the temperature dependence of our graphene based Hall transducers covering a major part of the temperature range which is typically specified for commercial devices, i.e. between room temperature and about 80 °C. For each temperature, the density dependence of the noise voltage spectral density (at one kHz) and the Hall coefficient were measured from which the magnetic field resolution can be calculated (see Eq. 5.27). A constant current between 20 and 30 μA was applied in order to ensure that the low frequency noise regime is dominated by flicker noise in case of our few micrometer sized devices. However, these current values are well below the threshold for current heating and the necessary bias voltage is low enough to not cause asymmetries due to the potential gradient.

Figure 5.11 shows the temperature dependence of two exemplary hBN-encapsulated graphene Hall transducers (each column corresponds to one device). The first device (left column) is of mediocre quality with lower mobility and larger contact resistances compared to the second device of highest quality (right column) which was fabricated following our proposed improvement steps involving thermal and AFM treatment (corresponding to the same device as in Fig. 5.10). The first device has Hall cross geometry and a larger footprint ($A = 7x3\,\mu m^2$) compared to the square-shaped second device ($A = 3x3\,\mu m^2$). Due to this difference in size, the obtained magnetic field resolutions are comparable even though the first device is of lower quality, i.e. lower mobility and larger residual total carrier density at the CNP which is, however, reflected in the smaller maximum Hall coefficient values.

As already discussed in Section 4.3.2 of the preceding chapter, graphene's resistivity away from the CNP increases with temperature mainly due to enhanced phonon scattering while the thermal activation of charge carriers does not play a role in this single carrier regime (see Figure 4.8). In addition to this rise in channel resistance, we also observe a slight but notable increase of the contact resistance away

from the CNP at higher temperatures. In total, the two terminal resistance between the source and drain contacts becomes larger as the temperature is increased. Since we operate our Hall transducers in constant current mode, this means that the bias voltage increases with temperature (at a fixed working point in terms of charge carrier density). Accordingly, the measured noise voltage spectral density, which is dominated by 1/f noise, grows as expected from the Hooge model: $N_{V,meas} \approx N_{V,1/f} \propto U \propto R(T) \cdot I$ (see Eq. 5.14 introduced in Section 5.3.1).

In order to distinguish between this effect and an intrinsic temperature dependence of the noise mechanism, we show the bias voltage normalized noise voltage spectral density (at one kHz), $N_{V,meas}/U$, in Figure 5.11. Within the studied temperature range[30] between room temperature and about 80 °C, we did not find a notable intrinsic temperature dependence of the 1/f noise. Hence, the observed increase of the noise voltage with temperature in current-mode operation can be attributed mostly to the temperature dependence of the input resistance and the bias voltage, respectively. Since the first device is of lower quality and exhibits a stronger R(T) dependence (not shown here explicitly, see Figure 4.8 in Section 4.3.2 for exemplary data), the change in 1/f noise with temperature is larger than for the second device. This is reflected in the B_{min} data in the third panel for which the absolute noise voltage spectral density is used for the calculation. The magnetic field resolution notably deteriorates in case of the first device, i.e. B_{min} increases over the entire density range. With higher mobility, different L/W ratio and lower contact resistances, the temperature dependence is much less pronounced for the second device (almost not distinguishable in the data point fluctuation).

While in constant current operation[31] the change in resistance with temperature governs the $N_V(T)$ and $B_{min}(T)$ dependencies in the single carrier regime away from the CNP, the thermal activation of charge carriers becomes important in the vicinity of the CNP. We already discussed the temperature dependence of the Hall coefficient in detail in Section 4.3 of the preceding chapter (see Figure 4.6 showing data from cryogenic to room temperature). The second row of panels in Figure 5.11 extend this temperature dependence investigation to higher temperatures. It

30 We also acquired data sets for 80 °C in case of the second device. However, this noise data contained some large spikes induced by the power controller of the heater stage used for this measurement (it was replaced in a later iteration of the setup which was then used for the first device). These spikes would overlay the other temperatures' curves. However, the general trend of no notable intrinsic temperature dependence also holds for this data set.

31 When in constant voltage operation, we expect the noise voltage to be almost temperature independent within the temperature range considered here, since the intrinsic temperature dependence is negligible. However, the current then decreases with increasing temperature because the resistance increases. Since $B_{min} = N_V/(R_H I)$ (Eq. 5.7), the magnetic field resolution again deteriorates with increasing temperature.

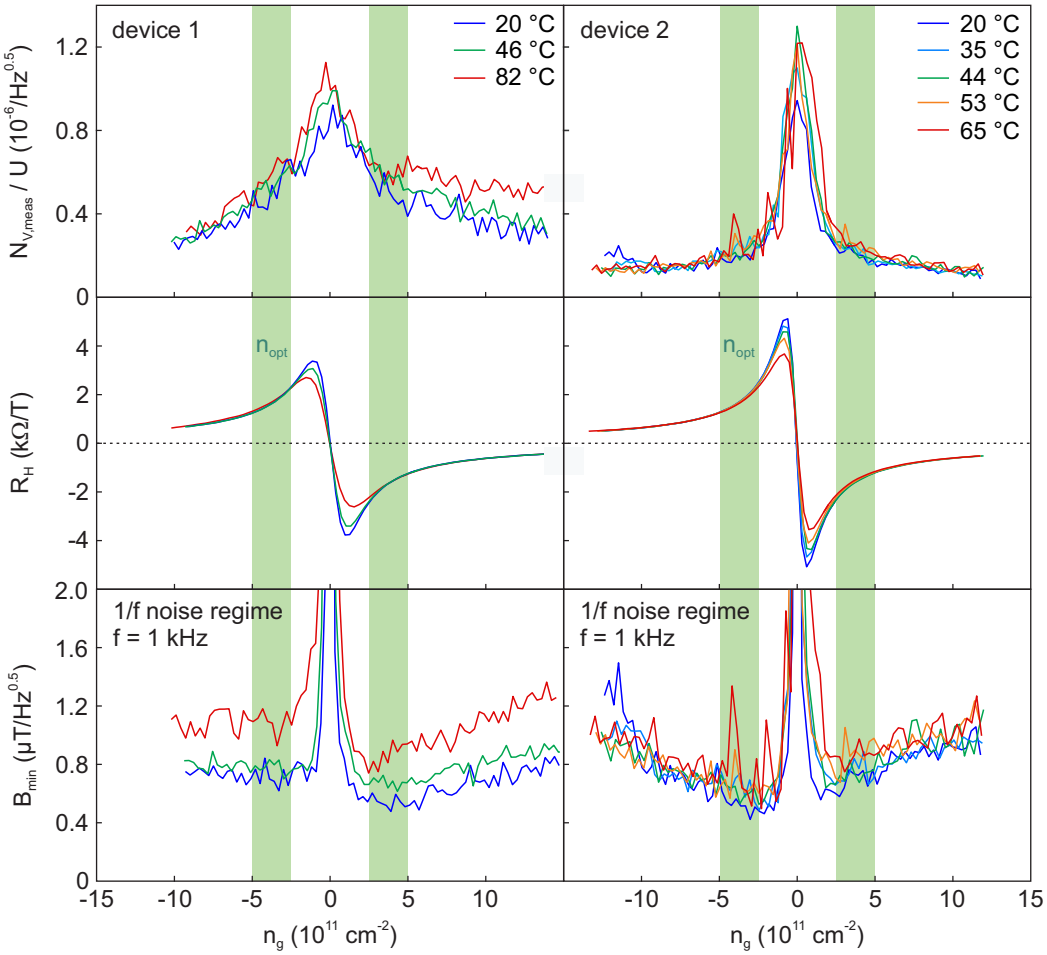

Figure 5.11. Temperature dependence of two hBN-encapsulated graphene Hall transducers within the typical temperature range of operation in commercial applications. Each column corresponds to one individual device and shows the density dependence of the normalized Hall voltage noise, the Hall coefficient and the magnetic field resolution for various temperatures. Noise voltages were obtained at a frequency of 1 kHz and the bias current was chosen such that 1/f noise dominated over the entire density range (i.e. $N_{V,meas} \approx N_{V,1/f}$). The device in the left column had lower mobility and higher contact resistances than the device of highest quality in the right column (which corresponds to the same device as in Fig. 5.10). The green areas mark the density range which we would recommend for operation ($n_{opt} \approx 2.5 - 5 \cdot 10^{11}\,cm^{-2}$).

was demonstrated that hBN-encapsulated devices at room temperature are mainly governed by thermal activation of charge carriers which outnumber the residual

carrier density due to extrinsic disorder. Accordingly, maximum Hall coefficient values between 4,000 and 5,000 V/AT are obtained which, however, are very sensitive to temperature. Moreover, the noise level increases at lower densities. Therefore, we recommend to operate hBN-encapsulated, high quality graphene transducers at higher densities, i.e. in the range of $n_{opt} \approx 2.5 - 5 \cdot 10^{11}\,cm^{-2}$ (green shaded areas in the figure). Due to the low extrinsic disorder and because the recommended density working point is about five times larger than the residual total carrier density at the CNP, $n_{t,0}(300K) \approx 6 - 7 \cdot 10^{10}\,cm^{-2}$, this corresponds to an operation in the single carrier regime. Hence, the Hall coefficient is almost temperature independent and the remaining temperature dependence of the magnetic field resolution can be attributed to $N_V(T)$ which is mostly determined by R(T) in the single carrier regime, i.e. $\rho_c(T)$ and $\rho(T)$. When the contact resistance is negligible, for example in large area hBN-encapsulated devices, which are not yet realized (see Section 5.8), we expect that the ultimate source of thermal drift is the change in mobility, i.e. phonon scattering. This is because the charge carrier density can be considered temperature independent at the recommended working point further away from the CNP where thermal activation of charge carriers plays no role (see panel d of Figure 4.8 in Section 4.3.2).

In the literature, only few reports provide noise data as a function of temperature. Ref. 404 and Ref. 409 studied the normalized noise power spectral density, N_{PSD}/U^2, in longitudinal configuration for hBN-encapsulated and SiO_2 supported graphene devices, respectively. They explored a temperature range between cryogenic temperatures and room temperature and found that the normalized noise power spectral density increases as the temperature rises. However, their data shows no longer a notable difference between 200 K and room temperature, i.e. the main increase in noise level occurs between cryogenic temperature and around 150 K. Hence, our results are consistent. Earlier works (Refs. 426 and 427) investigated the temperature dependence of the Hall voltage noise or B_{min} for graphene on SiO_2 devices between 77 K and 300 K. This means in transverse configuration as in our case. However, the provided data is not normalized by the bias voltage and hence does not account for the temperature coefficient of the resistance which makes a comparison difficult. In general, their data suggest an increase of the noise level and the B_{min} value when temperature is increased, in qualitative agreement with our measurements before normalization.

Thermal stability of the charge carrier density working point

In our measurements, we set a certain temperature and then tune the density via the Si substrate integrated back gate, that means using both the 290 nm thick SiO_2 and the bottom hBN layer as dielectric. After the data is acquired for each gate voltage step, the temperature is changed and the measurement protocol is repeated. For each temperature step and corresponding data set, the individual V_{CNP} value is determined and the single carrier density values for the axes in Figure 5.11 are calculated. This allows investigating the intrinsic temperature dependence of the Hall transducer's parameters without additional effects due to a potential $V_{CNP}(T)$ drift that is compensated in our analysis.

In practice, however, the Hall transducer is targeted to be operated at a certain density working point which is linked to a fixed gate voltage (at a certain temperature) that must be applied in order to tune the device to this ideally stable density working point. Even though hBN encapsulation of graphene reduces substantially both the initial doping and the gating hysteresis compared to uncovered devices on SiO_2, as demonstrated in Section 3.2.2 and in Figure 3.7, both values are still in the $10^{10} - 10^{11}\,cm^{-2}$ range. Moreover, as we already discussed in detail in Section 4.3 and Figure 4.6, there is a notable and sample-specific $V_{CNP}(T)$ drift. At higher temperatures above room temperature, we find this behavior to be more pronounced and shifts can correspond up to few $10^{11}\,cm^{-2}$, i.e. on the order of the recommended density working point. This means, that the temperature dependence of the parameters, for instance the Hall coefficient or magnetic field resolution, would become more complex assuming a fixed gate voltage operation. Effectively, the density working point then shifts slightly as a function of temperature.

Since we find the $n_g(V_g, T)$ relation to be sample-specific and since the main task of this work was to provide insight in the intrinsic temperature dependence of the Hall transducer's key performance indicators, we do not discuss here the temperature dependence for fixed gate voltage operation. This additional thermal instability is caused by extrinsic effects, mainly due to the use of the Si integrated back gate, the quite thick thermally grown SiO_2 dielectric and charged impurities at the SiO_2/hBN interface (see Sections 3.1.5, 3.2.2 and 5.8 for further discussion). Reducing the dielectric thickness, i.e. increasing the electric field strength, and avoiding the SiO_2/hBN interface improve the thermal stability of the working point. Furthermore, temperature coefficients must be determined on system level, combining the transducer element and the other electronic components. This also allows for additional compensation measures. In the last section of this chapter (Section 5.8), we

pick up this topic when discussing the remaining challenges for commercial graphene based Hall sensors.

5.7. Technical specifications, simulation and comparison

In this section, we provide a comprehensive summary of all important technical specifications and material parameters of high quality graphene Hall transducers which we obtained in this work. We then use these parameters in the models, whose validity we have proven in the preceding chapters, to simulate the potential performance of large-area, high quality graphene Hall elements with similar footprint as in today's commercial Hall sensors. Finally, our findings are compared to other studies on graphene Hall transducers as well as other material systems.

5.7.1. Key performance indicators and material parameters at a glance

Review of key performance indicators

In Chapter 4 and the previous two sections of this chapter, Sections 5.5 and 5.6, we have studied in detail the two main performance indicators of Hall transducers, the current-related sensitivity and the magnetic field resolution. Other key performance indicators are the offset at zero magnetic field and the non-linearity, as discussed in Section 5.2. The full-scale non-linearity, as defined in Eq. 5.9, and the offset resistance, $R_{off} = U_{off}/I$ (see Eq. 5.2), are shown in Figure 5.12 as function of charge carrier density for several temperatures. The data were measured on the same two hBN-encapsulated devices for which the Hall coefficient and the magnetic field resolution are given in Fig. 5.11 in Section 5.6.2.

For our few micrometer sized, hBN-encapsulated devices we find that the optimal density working point at room temperature, n_{opt}, is about $2.5 - 5 \cdot 10^{11}\,cm^{-2}$ (which equals to $4 - 5 \cdot n_{t,0}$). In this density range, the sensitivity is still reasonably high (20x compared to conventional Si Hall devices). The sensitivity's temperature dependence, however, is substantially smaller than at the density corresponding to maximum sensitivity (see Section 5.6.2). Moreover, the magnetic field resolution is optimal in the recommended range, i.e. B_{min} reaches a minimum (see Section 5.6.1). As discussed in detail in Sections 5.6.1 and 5.7.2, the increase of B_{min} with density

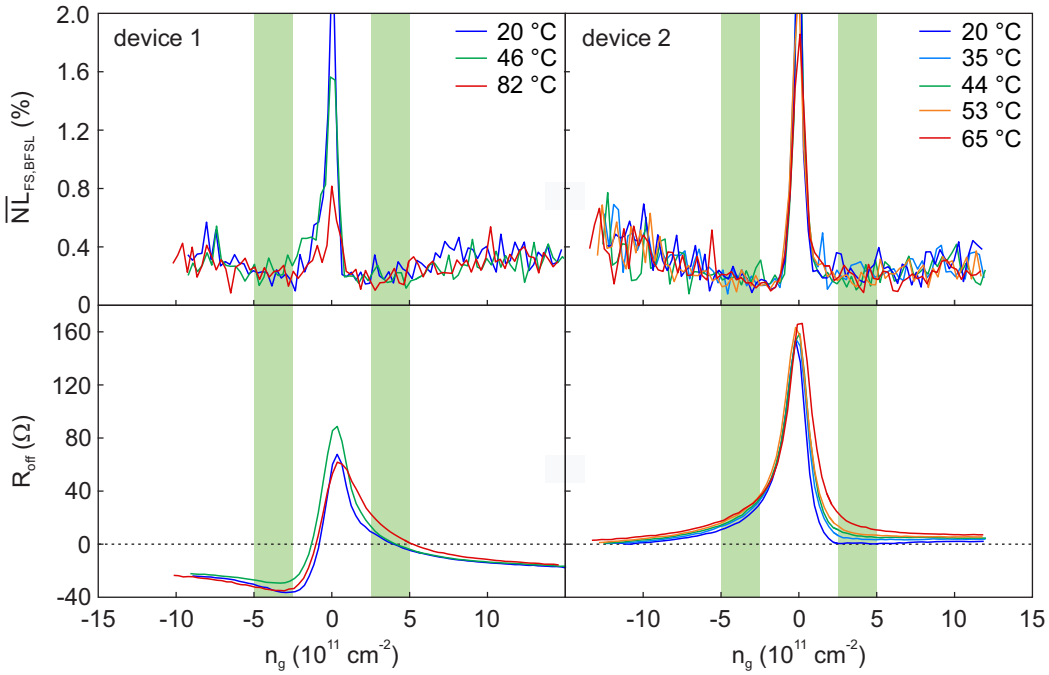

Figure 5.12. Density dependence of the mean non-linearity in % full-scale (BFSL method, ±50 mT range), as defined in Eq. 5.9, and the Hall offset resistance, $R_{off} = U_{off}/I$ (see Eq. 5.2). Data is shown for the same two hBN-encapsulated devices and the same different temperatures as in Fig. 5.11, each column corresponding to one device. Again, the green areas mark the density range which we recommend for operation (n_{opt}) with respect to sensitivity and magnetic field resolution in case of our few micrometer sized devices.

can be attributed to the comparably large contribution of the contact resistance to the total channel resistance in case of our few micrometer sized devices. For large area devices and/or negligible contact resistance, we expect that B_{min} continuously decreases towards higher densities (see following Section 5.7.2). Even then, however, the difference compared to the $2.5 - 5 \cdot 10^{11}\,\mathrm{cm}^{-2}$ range is marginal while the sensitivity decreases substantially and approaches ultimately values of Si Hall transducers (100 Ω/T). This is, for instance, the case for highly doped graphene on SiC (see Refs. [398, 418])

The mean (max) full-scale non-linearity in this recommended, quasi single carrier regime is $< 0.5\,\%$ ($< 1\,\%$) which is comparable to state-of-the-art Hall devices. We consider for our hBN-encapsulated devices a magnetic field range of ±50 mT such that $\mu^2 B^2 \ll 1$ holds. Otherwise, the Hall coefficient may show a magnetic field dependence in the two carrier regime in vicinity to the CNP which adds additional

non-linearity to the $U_H(B)$ trace, as shown in Fig. 4.1 and discussed in Section 4.1. However, this effect is already notably smaller in the above recommended density working point range. Good linearity is obtained over a much larger magnetic field range which is another advantage of sacrificing maximum sensitivity for a better general transducer performance. While the density dependence of the non-linearity is reproducible and shows weak to no temperature dependence in the investigated temperature range, the offset resistance as a function of density and temperature is subject to sample to sample variation. Two examples are shown in Figure 5.12. Some devices exhibit a resistive offset due to small alignment errors of the Hall leads, while for others the density dependence of the offset resembles the $R_H(n_g)$ curve (however, sometimes the sign is different). This could be caused by inhomogeneities in the channel. We expect that this variation decreases in case of an industrialized fabrication process of devices with larger area. Moreover, we emphasize that we consider here bare Hall transducers without additional on-chip components for further optimization. In today's Hall sensors, the spinning current method allows removing the static offset almost completely (see Section 5.2).

In Table 5.1 we have summarized the key performance indicators which we obtained on our few micrometer sized Hall transducers at room temperature. Temperature coefficients of these parameters are not provided here since one has to distinguish between intrinsic and extrinsic temperature effects, as already discussed above in Section 5.6.2. In practice, commercial devices will most likely be operated at a fixed gate voltage. Hence, the sample-specific thermal drift of V_{CNP} leads to additional complexity which we eliminated by considering the density dependence for each temperature (that means V_{CNP} was determined for each temperature). Then, only the intrinsic temperature dependence remains. Density dependencies for different temperatures are shown and discussed in Figure 5.12, Figure 5.11, Figure 4.8 in Section 4.3.2 and Figure 4.6 in Section 4.3. The offset showed in general a sample to sample variation which then also applies to its temperature dependence. Moreover, the noise voltage and, as a consequence, the magnetic field resolution showed different degrees of temperature dependence (which can likely be attributed to different contact resistances and device quality, see Section 5.6.2). Finally, the intrinsic temperature dependence of the Hall coefficient and channel resistivity are the two best understood, as discussed in Sections 4.3 and 4.3.2. They are governed by the temperature dependence of the residual carrier density at the CNP, i.e. by thermal activation of charge carriers, and the mobility (only for the resistivity). As discussed in Sections 4.1.2 and 4.3.1, the temperature dependence of the residual carrier density at the CNP, $n_{t,0}(T)$, can be described well for high quality, hBN-encapsulated devices using Eq. 4.4 with a Fermi velocity of $1.6 - 1.7 \cdot 10^6\,\mathrm{m/s}$. Furthermore, temperature coefficients should be assessed on system level, taking into account both

Key performance indicator	Definition	hBN-encapsulated	on SiO_2
Sensitivity (current related)	$S_I = R_H$ (Eq. 5.1)	$S_{I,max}$ between 4,000 and 6,000 Ω/T (Fig. 4.2); ≈ 2,000 Ω/T at n_{opt} (much smaller T-drift)	$S_{I,max}$ between 1,500 and 2,000 Ω/T (Fig. 4.2)
Magnetic field resolution	$B_{min,sat}$ in the 1/f noise regime (Eqs. 5.7 and 5.19)	At n_{opt} and f = 1 kHz: several hundred nT/$\sqrt{\text{Hz}}$ (for channel width of few μm); ≈ 1 − 10 nT/$\sqrt{\text{Hz}}$ for large area devices (Section 5.7.2); $B_{min} \cdot W \approx$ 500 − 1,500 $\frac{\text{nT}\,\mu\text{m}}{\sqrt{\text{Hz}}}$.	$B_{min} \cdot W \approx$ several 1,000 to 20,000 $\frac{\text{nT}\,\mu\text{m}}{\sqrt{\text{Hz}}}$ (see Table 5.3).
Hall offset resistance	$R_{off} = U_{off}/I$ (Eq. 5.2)	sample dependent, for exemplary $R_{off}(n_g)$ dependence see Fig. 5.12; ≈ 20 − 40 Ω in recommended density range	similar as hBN
Non-linearity (full-scale)	$NL_{FS,BFSL}$ (Eq. 5.9)	mean < 0.5%, max < 1% for ±50 mT (Figs. 5.12 and 4.1)	similar as hBN

Table 5.1. Summary of all KPIs of graphene based Hall transducers at room temperature, without optimization through additional on-chip components. Measurements were conducted on few micrometer sized, exfoliated samples (in case of SiO_2 also some CVD-grown devices with larger dimensions were considered). In case of hBN-encapsulation, an increase in dimensions might lead to further improvement of the magnetic field resolution due to a potential reduction in contact resistance so that its contribution to the overall resistance becomes negligible (see Section 5.7.2).

the transducer and the other electronic on-chip components. This also allows for additional temperature drift compensation measures.

Moreover, we did not specify the power or current consumption in Table 5.1. For constant current operation, the power consumption scales quadratically with the

applied current, $P_{el} = R\,I^2$ (see Eq. 5.1). The input resistance R_{in}, i.e. the two terminal resistance between source an drain, is a function of channel resistivity, device geometry (i.e. L/W ratio) and contact resistance. We found the contact resistance to vary among the fabricated devices as well as to contribute approximately equally to the input resistance in case of few micrometer sized devices. Assuming a mobility of $70,000\,\mathrm{cm}^2/\mathrm{Vs}$, the channel resistivity is about $300\,\Omega$ at the recommended density working point of $\approx 3 \cdot 10^{11}\,\mathrm{cm}^{-2}$. For a Hall cross geometry (L/W ≈ 3), the channel resistance is $900\,\Omega$. For such devices with few micrometer width, about the same value can be expected for the contact resistance. In our measurements on few micrometer sized devices, $B_{min,sat}$ is reached for a current larger than few ten μA. Assuming $I = 30\,\mu\mathrm{A}$ and $R_{in} \approx 2\,\mathrm{k}\Omega$, the power dissipation in the transducer amounts to 60 mW. When scaling the device dimensions to today's typical Hall transducer dimensions in commercial sensors, the channel resistance remains the same, while the contribution of the contact to the total two terminal input resistance becomes negligible. However, as discussed in detail in Section 5.7.2, $B_{min,sat}$ is reached for an about one order of magnitude larger current, i.e. few hundred μA. For $300\,\mu$A, the power consumption would be about 270 mW. Finally, the transducer of the Hall sensor should not be assessed without taking into account the other on-chip components which also consume power but allow for reducing the dissipation in the transducer part. For instance, the current can be applied only during a short period of time, i.e. in form of square wave pulses. This reduces the effective current on time average while a strong signal can be extracted by means of the boxcar averaging technique. Finally, in battery driven applications, the above consideration of power consumption is less relevant. Instead, the current consumption constitutes the important metric, given the battery's capacity in mAh. The available supply voltage of about 4 V for Li-ion batteries is more than enough to drive the Hall transducer (rather, it must be reduced such that only few hundred μA are applied as maximum). As discussed in Section 5.7.2, large area graphene Hall transducers are expected to reach $B_{min,sat}$ when operated in the $100 - 200\,\mu$A current range (without measures like boxcar averaging). This is comparable to the current consumption typically specified in existing commercial devices.

Key material parameters for simulation of graphene Hall elements

Besides the above list of the key performance indicators, we find it important and helpful to review the key material parameters and models that allow us to predict the Hall element performance in terms of sensitivity, magnetic field resolution, input resistance and power consumption for any geometry, dimensions as well as operating conditions (such as current, density or frequency). Table 5.2 provides a summary

of these models and material parameters that we obtained on our few micrometer sized devices.

The channel approximation (Eq. 2.30) introduces the ad-hoc *residual total charge carrier density at the CNP* to approximate the total charge carrier density as a function of gate controlled single carrier density, n_g. This approximation replaces then the single carrier density in the Drude models for the resistivity (Eq. 3.1) and Hall coefficient (Eq. 4.3). Moreover, we can use it in the Hooge model to consider the two carrier case by replacing n_t with $\sqrt{n_{t,0}^2 + n_g^2}$ in Eq. 5.14 in Section 5.3.1. Since the two-terminal input resistance depends on the resistivity and the bias voltage depends on the input resistance for constant current operation, $n_{t,0}$ also has an indirect impact on the 1/f noise and thermal noise floor (see Eq. 5.10).

Apart from the residual total carrier density, the *mobility* determines the resistivity (Eq. 3.1). Accordingly, the same chain of direct and indirect dependencies as above applies. Moreover, the mobility is a measure of electronic quality (extrinsic disorder). As discussed in Section 5.5.3, devices with higher quality, i.e. mobility, exhibit a lower Hooge parameter.

The dimensionless *Hooge parameter* determines the 1/f noise level in the empirical Hooge model (Eq. 5.13). Therefore, it dictates for which critical current and corner frequency (I_c and f_c, respectively) the transition from the thermal noise floor to the 1/f noise regime occurs and the magnetic field resolution eventually saturates (given a certain channel area und density). We again note that the Hooge parameters provided here were obtained on few micrometer sized devices which exhibit a relatively large contribution of contact resistance. Hence, the given values can be considered as upper, conservative boundary, as already discussed at the end of Section 5.5.3.

Finally, a parameter we came across several times in Sections 5.3.3, 5.5 and 5.6.1, is the *contact resistivity*, ρ_c (in units of $\Omega\,\mu$m). It defines the contact resistance $R_c = 2\rho_c/W$ (Eq. 5.11). For few micrometer sized devices, it contributes notably to the input resistance, comparable to the channel resistance in case of high quality, hBN-encapsulated devices with high mobility. Hence, it has an impact on the power consumption and both the thermal noise floor as well the 1/f noise in case of constant current operation (see Eqs. 5.10 and 5.14). Other studies also found that 1/f noise in exfoliated, few micrometer sized devices is governed to a notable extent by resistance fluctuations at the contacts [412, 413, 425]. We find the contact resistivity to be sample as well as density dependent. In general, increasing the density reduces the contact resistivity. However, also the dependence on the density varies among different samples and thus no exact functional dependency can be given. For our

Key material parameter	Models and dependencies	hBN-encapsulated	on SiO_2
residual total carrier density at the CNP ($n_{t,0}$)	channel approximation: $n_t = \sqrt{n_{t,0}^2 + n_g^2}$ (Eq. 2.30) $\frac{1}{\rho} = e\,\mu\sqrt{n_{t,0}^2 + n_g^2}$ (Eq. 3.1) $R_H = -\frac{n_g}{e\,(n_{t,0}^2+n_g^2)}$ (Eq. 4.3) N_V (Eqs. 5.10, 5.14)	$\approx 7 \cdot 10^{10}\,cm^{-2}$ (Fig. 3.8 and Section 4.3.1)	between 10 and $30 \cdot 10^{10}\,cm^{-2}$ (Fig. 3.8 and Section 4.3.1); for the calculation in Section 5.7.2 we assume $15 \cdot 10^{10}\,cm^{-2}$ (optimistic)
mobility (μ)	$\rho(\mu) \rightarrow R(\rho) \rightarrow U \rightarrow N_V$ (Eqs. 5.10, 5.14); measure of device quality, $\rightarrow \alpha_H(\mu)$ (Section 5.5.3)	density dependent (Figs. 3.8 and 4.3), $\approx 70,000\,cm^2/Vs$ for few $10^{11}\,cm^{-2}$	$\approx 7,000\,cm^2/Vs$ (Fig. 3.8)
Hooge parameter (α_H)	1/f noise, $N_{1/f}$ (Eq. 5.14)	$\approx 10^{-5}$ (Section 5.5.3, Fig. 5.8)	$\approx 10^{-3}$ (Section 5.5.3, Fig. 5.8)
contact resistivity (ρ_c)	$R_c = 2\rho_c/W \rightarrow R \rightarrow U \rightarrow N_V$ (Eqs. 5.10, 5.14)	sample dependent, typically from 1,000 to $3,000\,\Omega\,\mu m$ for $10^{11} - 10^{12}\,cm^{-2}$ (see main text)	Typically smaller than for side contacts in van der Waals heterostructures. In Section 5.7.2 we assume $3,000\,\Omega\,\mu m$ as conservative, upper boundary for both hBN and SiO_2.

Table 5.2. Key material parameters at room temperature and models to calculate the performance of graphene based Hall transducers in terms of sensitivity, magnetic field resolution and power consumption (see Section 5.7.2).

processing routine involving a short oxygen plasma treatment before evaporation of Cr-Au contacts (with re-exposure to air between the two steps in separate machines), we typically obtain 1,000 to $3,000\,\Omega\,\mu m$ for edge contacts in the $10^{11} - 10^{12}\,cm^{-2}$ density range. For areal contacts in case of graphene on SiO_2, i.e. for contacts on top of the graphene surface, these values are typically lower (but not substantially). However, due to the lower mobility for devices on SiO_2, the relative contribution

of the contacts to the input resistance is much smaller than in the case of hBN-encapsulated, high mobility devices. Another study (Ref. 10) reports contact resistivities between 200 and 1000 $\Omega\,\mu$m for the above density range and edge contacts. Obviously, the contact resistances vary depending on the used etching and evaporation equipment and recipes. For instance, we used similar recipes as in Ref. 10 but a different dry etching machine. This leaves room for further improvement of graphene Hall transducers, i.e. by optimizing contact resistances. Some studies investigated how the contact resistivity can be reduced [10, 414, 415, 428, 429]. For edge contacts in van der Waals heterostructures, the contact resistivity is strongly influenced by the functionalization of the graphene edge, e.g. through plasma treatment before metal evaporation, and the choice of metal combination as contact material [10]. For areal contacts, it helps to increase the effective peripheral length through contact patterning [428]. Finally, multilayer graphene can be used to contact the monolayer graphene sheet for which the lowest contact resistances are achieved in case of an optimal alignment angle between the lattices [415]. It is, however, doubtful that the latter approach can be implemented in large scale production due to the increased complexity and need for additional transfers of multilayer graphene. To conclude, it will be very challenging and probably impossible to reach the low contact resistivity that is obtained for bulk material systems with mature fabrication technology such as Si ($50\,\Omega\,\mu$m, Ref. 428).

Testbed: Simulation of micrometer sized Hall elements

We can now utilize the parameters and formulas given in Table 5.2 to calculate the sensitivity, the noise voltage spectral density and the magnetic field resolution for an arbitrary graphene Hall transducer. At first, we assume a few micrometer sized hBN-encapsulated Hall cross device with $\mathrm{L} = 9\,\mu$m and $\mathrm{W} = 3\,\mu$m which is comparable to our samples fabricated from exfoliated flakes. Since $\mathrm{L/W} = 3$, the geometric factor for the Hall coefficient is approximately unity and neglected. Moreover, we assume an operation at $\mathrm{f} = 1\,\mathrm{kHz}$. We emphasize that we make two simplifications in the following simulation of graphene Hall elements. First, a density independent mobility is used ($\mu = 70,000\,\mathrm{cm}^2/\mathrm{Vs}$ in case of hBN-encapsulation). This is, however, a good approximation in the density range of interest, i.e. $10^{11} - 10^{12}\,\mathrm{cm}^{-2}$, as shown in Fig. 4.3). Second, the contact resistivity is also assumed to be density independent and a conservative, high value of $\rho_\mathrm{c} = 3,000\,\Omega\,\mu$m is used in the computation (if not stated otherwise). By doing so, we underestimate the channel resistance at higher densities since the mobility is decreasing at higher densities in real devices which is not taken into account. But, we also overestimate the contact resistance at higher densities since a quite high, density independent value is assumed while ρ_c

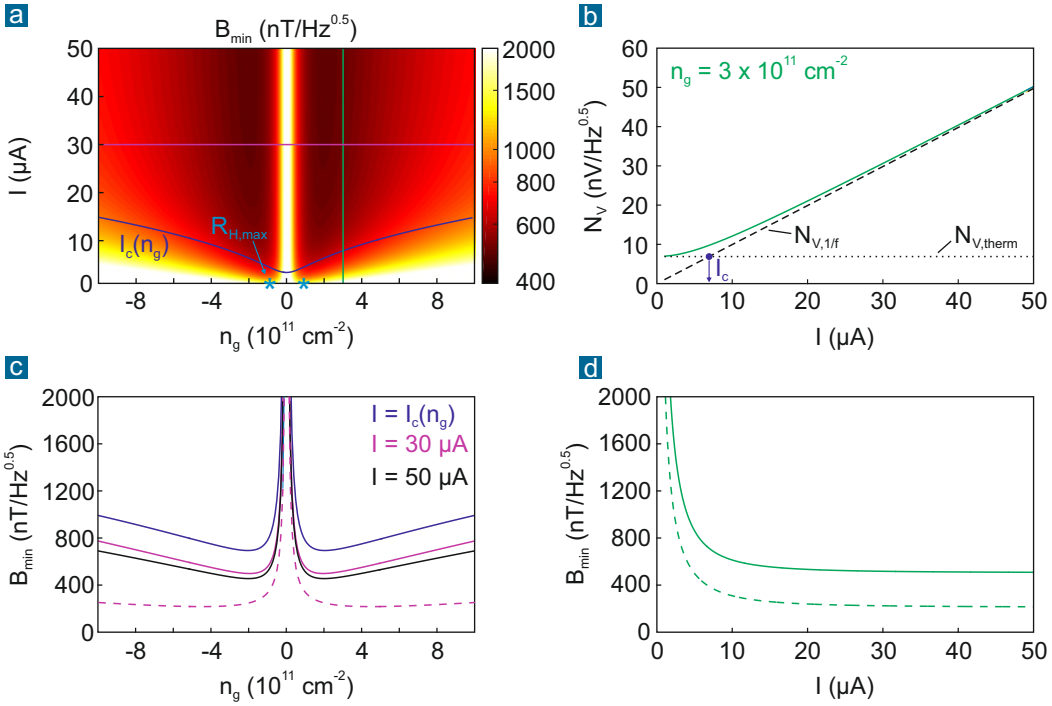

Figure 5.13. Calculation of the magnetic field resolution as function of applied current and density for a hBN-encapsulated graphene Hall transducer at f = 1 kHz, assuming Hall cross geometry and few micrometer dimensions (L = 9 μm, W = 3 μm). This corresponds to typical devices that can be fabricated today and were measured in this work. The two carrier models and the parameters in case of hBN-encapsulation as summarized in Table 5.2 were used. A conservative, high contact resistivity of $\rho_c = 3,000\,\Omega\,\mu$m is assumed (which we found as upper boundary for our devices in the mid to high density range). For simplification, no further density dependence of ρ_c was taken into account. The dashed lines in panel c and d correspond to a contact resistivity of $\rho_c = 500\,\Omega\,\mu$m. See main text for details.

is found to decrease with increasing density in real devices (even though the exact functional dependence varies among different devices). Hence, both simplifications compensate each other to some extent. The remaining materials parameters are $n_{t,0} = 7 \cdot 10^{10}\,\mathrm{cm}^{-2}$ and $\alpha_H = 10^{-5}$, as listed in Table 5.2.

Figure 5.13 shows the simulation result for the above specified few micrometer sized Hall element. In panel a, B_{min} is computed as function of applied current and gate controlled single carrier density. The blue curve corresponds to the density dependent, critical current value for which the 1/f noise equals the thermal noise floor (as shown in panel b). In analogy to the corner frequency, f_c, we use here the notation

I_c. I_c increases slowly with larger densities due to the comparatively large contribution of the contact resistance in case of the few micrometer sized device considered here (see the following section for a detailed discussion). For fixed density, 1/f noise quickly dominates and B_{min} saturates for currents larger than I_c, as demonstrated in panel b and d for $n_g = 3 \cdot 10^{11}\,cm^{-2}$ (corresponding to the green vertical line in panel a). This density value was chosen since it lies within the density range that we found to be optimal in our measurements on few micrometer sized samples ($n_{opt} \approx 2.5 - 5 \cdot 10^{11}\,cm^{-2}$), as discussed in the preceding sections. Also, our simulation suggests that the best magnetic field resolution can be obtained approximately in this range[32] but not at the density working point for which maximum sensitivity is obtained, i.e. $R_{H,max}$. The latter is marked by the turquoise stars in panel a of Figure 5.13.

Moreover, the simulation reveals that we reached the regime of prevailing 1/f noise over the entire density range that we studied in our experiments ($\pm 1 \cdot 10^{12}\,cm^{-2}$) since we typically applied currents stepwise from 10 to 30 μA (for devices with $> 5\mu$m width even up to 50 μA). Panel c of Figure 5.13 shows that B_{min} saturates over the density range. Almost no more improvement in magnetic field resolution is achieved when the current is increased from 30 μA (corresponding to the magenta horizontal line in panel a) to 50 μA.

To summarize, our simulation is based on only four material parameters and the small set of models and approximations (empirical Hooge model, channel approximation and two carrier Drude models for resistivity and Hall coefficient) listed in Table 5.2. However, it reproduces qualitatively the density and current dependence of the sensitivity and magnetic field resolution that we measured on our few micrometer sized devices (compare the result of the simulation in Fig. 5.13 to the measurements in Fig. 5.10 in Section 5.6.1). Also the frequency dependence of our data is captured well by the Hooge model for 1/f noise, as demonstrated earlier in Section 5.5.2. Hence, we are confident that our simulation model for graphene Hall transducers allows us also to predict properly the performance of devices with larger footprint, as discussed in the following section.

Here, for few micrometer sized devices, the simulation captures nicely the minimum of B_{min} within an optimal density range (B_{min} then increases at higher densities). This minimum in the $B_{min}(n_g)$ curve is more pronounced in case of a high contact resistivity and few micrometer dimensions. Besides our standard value for ρ_c in the

32 The lowest $B_{min,sat}$ is obtained at $n_g = 1.8 \cdot 10^{11}\,cm^{-2}$ in the simulation. Considering that several models for the Hall transducer simulation are empirical or use approximations as well as taking the two simplifications we made into account, this computational result agrees well with the experiments.

simulations ($3000\,\Omega\,\mu$m), we also ran a calculation with $500\,\Omega\,\mu$m corresponding to the dashed lines in panels c and d of Figure 5.13. In this case, the minimum in the $B_{min}(n_g)$ curve becomes negligible for a fixed current of $30\,\mu$A. However, also the $I_c(n_g)$ curves become steeper (not shown), reaching about $30\,\mu$A instead of $15\,\mu$A in the high density regime. This means that for the dashed line in panel c, the B_{min} values at high densities have not yet reached (full) saturation. Applying larger currents than $30\,\mu$A leads to (slightly) lower saturation values at high density than shown in panel c. Considering only the magnetic field resolution, there is no more an optimal density range ($n_{opt} \approx 2.5 - 5 \cdot 10^{11}\,\mathrm{cm}^{-2}$). However, working at high densities substantially reduces the current related sensitivity while the magnetic field resolution, i.e. effectively the signal to noise ratio, is only slightly improved. Moreover, it is necessary to apply larger currents to obtain these slightly lower values. This is critical in battery driven applications where current consumption is even more important than power consumption. Realistically, an upper limit for the current will be specified for which B_{min} should be minimized, having also certain minimum requirements for the sensitivity. Fortunately, graphene Hall elements show great flexibility (due to the density tunability) to meet all these opposing requirements.

Conclusion: For few micrometer sized devices and $\rho_c = 3000\,\Omega\,\mu$m, $B_{min,sat}$ reaches a minimum of about $500\,\mathrm{nT}/\sqrt{\mathrm{Hz}}$ in the few $10^{11}\,\mathrm{cm}^{-2}$ density range. For a smaller contact resistivity of $\rho_c = 500\,\Omega\,\mu$m, $B_{min,sat}$ is about $200\,\mathrm{nT}/\sqrt{\mathrm{Hz}}$ over a much wider density range. At higher densities a further reduction of this value is possible (by 10-20 %), however at the cost of a substantially lower sensitivity and higher current consumption. As will be discussed in detail in Section 5.7.3, the parameter $B_{min} \cdot W$ can be provided in order to compare different material systems for Hall elements. However, the definition of these parameters makes strictly only sense when the contact resistances are negligible (see Sections 5.3.2 and 5.7.3). This is not valid in the case of our few micrometer sized devices with high mobility. Ignoring this, both our simulation as well our experiments find $B_{min} \cdot W$ to range between $500\,\mathrm{nT}\,\mu\mathrm{m}/\sqrt{\mathrm{Hz}}$ and $1500\,\mathrm{nT}\,\mu\mathrm{m}/\sqrt{\mathrm{Hz}}$ at a frequency of 1 kHz. Finally, we note that our already well performing simulation model could be improved even further by taking into account a functional dependence of the contact resistivity and the mobility on the density (i.e. the exact theory of the dominating scattering mechanisms).

5.7.2. Simulation of large-area graphene Hall elements

High quality, hBN-encapsulated graphene Hall transducers

We can now employ the above introduced and tested simulation model and the key material parameter values in Table 5.2 to estimate the room temperature performance of transducers with a larger footprint that is comparable to today's commercial devices, i.e. in the one hundred micrometer range. All computations are performed for $f = 1\,kHz$. Like for the simulation of few micrometer sized devices, B_{min} is calculated as a function of current and gate controlled single carrier density assuming a Hall cross geometry with $L = 150\,\mu m$ and $W = 50\,\mu m$ ($L/W = 3$). Figure 5.14 shows the results for hBN-encapsulated graphene (panel a, b) as well as for graphene on SiO_2 (panel c, d). In both cases, a density independent contact resistivity of $\rho_c = 3000\,\Omega\,\mu m$ is assumed and the other parameters are used according to Table 5.2. In our experiments on few micrometer sized devices we found that a current density of $10\,\mu A/\mu m$ does not cause any asymmetries or a shift of V_{CNP}. Accordingly, we expect that devices with such large footprint can sustain currents up to $500\,\mu A$. Considering the strict requirements in terms of current consumption in case of battery powered applications, the current rating will likely be restricted to $100 - 200\,\mu A$ (however, larger currents could still be applied in short pulses when using the boxcar averaging method).

Comparing the large area with the few micrometer sized, hBN-encapsulated devices (panel a, b of Figure 5.14 and Figure 5.13, respectively), it is obvious that the $I_c(n_g)$ curve does not flatten out towards high density. This means that large currents must be applied in order to reach the saturation value of the magnetic field resolution (the higher the density, the larger the current). This is due to the comparatively smaller contact resistance which scales with $1/W$ while the channel resistance remains the same as for the few micrometer sized devices (same mobility and geometry). In contrast to the large area graphene on SiO_2 devices discussed below, there exists always a density working point, given a certain current, for which B_{min} is minimal. Increasing the density beyond this point leads again to an increase of the magnetic field resolution. However, this optimal density working point shifts to higher densities for larger currents. Unambiguously, the density for maximum sensitivity, i.e. $R_{H,max}$, is not optimal in terms of magnetic field resolution. For a maximum current between 100 and $200\,\mu A$, our simulation suggests that the lowest B_{min} values are $10 - 20\,nT/\sqrt{Hz}$ at about $3 \cdot 10^{11}\,cm^{-2}$ (crossing of the magenta and green line in panel a of Fig. 5.14). As shown in panel b, increasing the current further does no longer improve the magnetic field resolution drastically at this density. The lowest

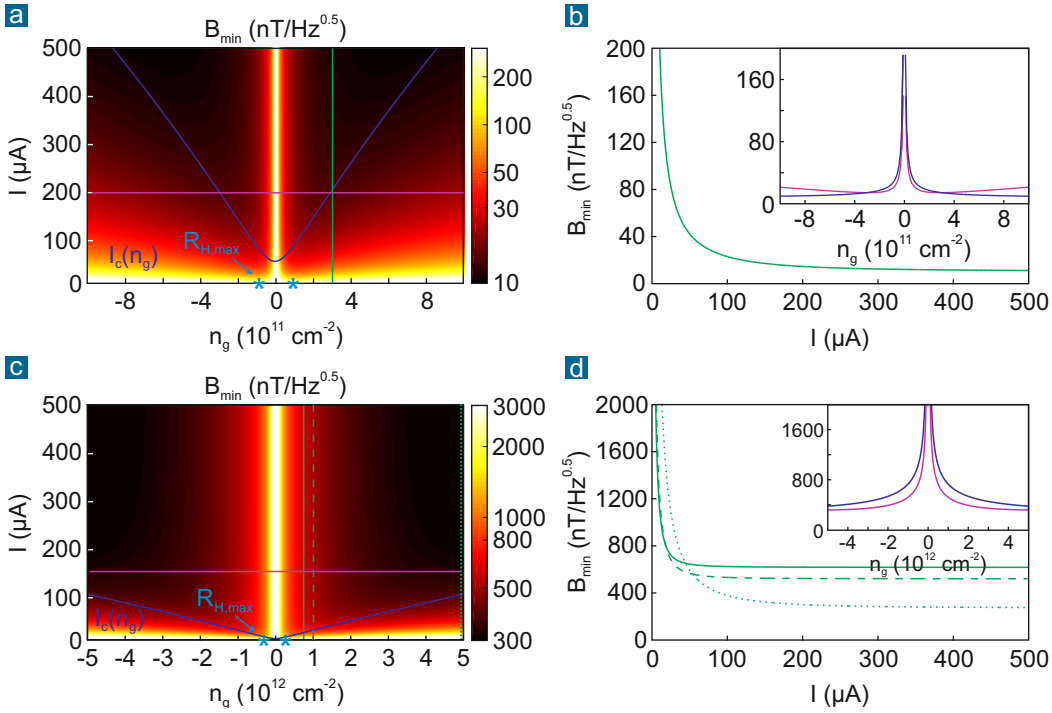

Figure 5.14. Calculation of the magnetic field resolution as function of applied current and density for large-area graphene Hall transducers at f = 1 kHz. Panel a, b: hBN-encapsulated Hall cross with L = 150 μm and W = 50 μm and $\rho_c = 3,000\,\Omega\,\mu$m. All other parameters as given in Table 5.2 for hBN-encapsulation. Panel c, d: Graphene on SiO_2 Hall cross with same dimensions and $\rho_c = 3,000\,\Omega\,\mu$m (which is, however, negligible compared to the channel resistance). Again, all other parameters are listed in Table 5.2. Note the different current range compared to Fig. 5.13, as well as the different scaling of the B_{min} and n_g axes between hBN (panel a,b) and SiO_2 (panel c,d). See main text for detailed discussion.

$B_{min,sat}$ values within the entire parameter space of density and current, as considered in panel a, are found to be in the $5 - 10\,\mathrm{nT}/\sqrt{\mathrm{Hz}}$ range. However, achieving this minor improvement would mean a substantially lower sensitivity and increase of the current consumption. We also performed a calculation assuming $\rho_c = 500\,\Omega\,\mu$m and found again only marginal improvements of the magnetic field resolution (not shown). Moreover, another simulation was run for a square geometry (150x150 μm^2) and even higher currents of up to 800 μA were considered. In this case, the lowest $B_{min,sat}$ values reach about $3\,\mathrm{nT}/\sqrt{\mathrm{Hz}}$.

To conclude, we predict that hBN-encapsulated Hall transducers with comparable footprint and maximum current ratings than in today's commercial devices can realistically[33] reach a magnetic field resolution in the $1 - 10\,\mathrm{nT}/\sqrt{\mathrm{Hz}}$ range at $\mathrm{f} = 1\,\mathrm{kHz}$ when the density working point is optimized solely with respect to the resolution. This corresponds to the lower end of the $\mathrm{B_{min}} \cdot \mathrm{W}$ range (between $500\,\mathrm{nT}\,\mu\mathrm{m}/\sqrt{\mathrm{Hz}}$ and $1500\,\mathrm{nT}\,\mu\mathrm{m}/\sqrt{\mathrm{Hz}}$) which was stated above for few micrometer sized Hall elements. This is due to the smaller contribution of the contact resistance in such large-area devices. The more the contact resistance can be reduced, the better magnetic field resolutions can be achieved, however, requiring a higher density working point and larger currents (which can be in conflict with other requirements such as high sensitivity and low current consumption).

Graphene Hall transducers on SiO_2

For comparison, we also conducted a simulation of large area graphene Hall transducers on SiO_2. Same geometry and dimensions ($\mathrm{L} = 150\,\mu\mathrm{m}$ and $\mathrm{W} = 50\,\mu\mathrm{m}$) as for the hBN-encapsulated device and again a frequency of 1 kHz were assumed. With CVD grown graphene or graphene on SiC, having comparable material parameters as on SiO_2, these dimensions can already be realized today (see Refs. 242, 348, 398, 418). We use the optimistic material parameters for graphene on SiO_2 given in Table 5.2. Since the mobility is one order of magnitude lower than in the case of hBN-encapsulation ($7,000\,\mathrm{cm}^2/\mathrm{Vs}$), the contribution of the channel to the total resistance prevails over the entire density range. As a consequence, the actual assumed value of the contact resistivity (here $3,000\,\Omega\,\mu\mathrm{m}$) is not pivotal. We found no substantial change in the simulation results when another, i.e. lower, value is assumed. Moreover, a rather low, i.e. optimistic, residual total charge carrier density at the CNP of $15 \cdot 10^{10}\,\mathrm{cm}^{-2}$ and a two orders of magnitude higher Hooge parameter than for hBN-encapsulated graphene ($\alpha_\mathrm{H} = 10^{-3}$, see Fig. 5.8) complete the input for the simulation.

Panel c and d of Figure 5.14 show the result of the simulation. Note the different range of the charge carrier density which is five times larger compared to the hBN-encapsulated graphene case in panel a and b, respectively. Since the 1/f noise is substantially larger in case of graphene on SiO_2 (α_H is two orders of magnitude larger),

33 As discussed in Section 5.5.3, the higher contribution of the contact resistance in our measurements on few micrometer sized devices may result in an overestimation of the Hooge parameter. Scaling the device area could to some extent also entail a further reduction of α_H itself, that means beyond the already considered reduction of the contact resistance in the noise formula (Eq. 5.14). Hence, our simulation results can be seen as a conservative, upper bound.

the $I_c(n_g)$ curve is less steep than for large-area, hBN-encapsulated graphene. This means that 1/f noise already dominates over the entire density range for currents above $150 - 200\,\mu A$ and B_{min} saturates. Moreover, since the contact resistance is negligible compared to the channel resistance in case of the large area device on SiO_2, $B_{min,sat}$ decreases continuously with increasing density (as expected from Eq. 5.20 in Section 5.3.2). Hence, there is no optimal density working point. Instead, the higher the density the better the magnetic field resolution becomes, in particular if larger currents may be applied in compliance to current consumption restrictions. However, as already discussed above, the sensitivity is substantially lower for densities in the $10^{12}\,cm^{-2}$ range. Even when operating at high densities, our simulation finds that the best obtainable B_{min} values for large area graphene Hall elements on SiO_2 are limited to the $100 - 400\,nT/\sqrt{Hz}$ range at $f = 1\,kHz$ (see panel d of Fig. 5.14). This is one to two orders of magnitude higher, i.e. worse, than for hBN-encapsulated graphene with the same footprint. This shows the great potential of large area, hBN-encapsulated graphene devices, if they can be realized in the future. Finally, the width normalized magnetic field resolution, $B_{min} \cdot W$, ranges between $5,000\,nT\,\mu m/\sqrt{Hz}$ and $20,000\,nT\,\mu m/\sqrt{Hz}$.

The findings of our simulation of the magnetic field resolution agree well with experimental data of large area graphene Hall transducers on SiO_2 or SiC (see Section 5.7.3 and for instance Refs. 348, 398). In some cases, the rather high initial doping of graphene on these substrates is exploited intentionally without the need for a back gate. Instead, the initial doping already results in a high density working point which we found to be optimal for graphene with lower mobility. For instance, Ciuk and coworkers (Ref. 398) studied graphene on SiC Hall elements with dimensions of $L = 500\,\mu m$ and $W = 150\,\mu m$. They reported a sensitivity of $87\,\Omega/T$ which corresponds to a density of $7 \cdot 10^{12}\,cm^{-2}$. Running a simulation using our graphene on SiO_2 parameters, which are very similar to the SiC case, the simulation predicts $B_{min,sat}$ of about $100\,nT/\sqrt{Hz}$ at $f = 1\,kHz$. This agrees perfectly with the experimental findings of Ciuk and coworkers. Moreover, the specifications sheet of the first commercially available[34] graphene Hall element with analog operation, i.e. without any additional on-chip optimization, reports a sensitivity of $150\,\Omega/T$ and a resolution of 20 ppm with reference to 1 T at 1 Hz. This translates to a density working point of about $4 \cdot 10^{12}\,cm^{-2}$ and resolution of $\approx 600\,nT/\sqrt{Hz}$ at 1 kHz for the specified nominal supply current of $200\,\mu A$. Again, this is in good agreement with our findings.

34 Publicly available data sheet of the graphene Hall sensor GHS01AT on the website of Cambridge based start-up Paragraf (www.paragraf.com).

Conclusion

Likely the most important property of graphene based systems is that graphene's density can be varied over a wide density range, e.g. by means of a gate. This provides the possibility to adapt the working point of graphene Hall elements to application specific requirements, e.g. in terms of sensitivity, resolution, linearity and temperature coefficients. When assuming that all graphene Hall devices on different substrates are operated at their (specific) optimal density working point in terms of best magnetic field resolution, we find that few micrometer sized, hBN-encapsulated graphene devices reach similar $B_{min,sat}$ values as large area graphene on SiO_2 devices. If few tens of nanometer thick hBN films could be grown on large scale, hBN-encapsulated graphene Hall elements with a footprint comparable to today's commercial devices would even outperform their counterparts on SiO_2 by one or two orders of magnitude. The best magnetic field resolution for hBN-encapsulated devices is obtained at densities for which the sensitivity is still at least one order of magnitude larger than that of conventional Si Hall devices ($\approx 100\,\Omega/\mathrm{T}$). However, for graphene Hall elements on SiO_2 a higher density working point is favorable to improve the magnetic field resolution. This results in low sensitivity values that are comparable to Si Hall sensors. However, working at such high densities also improves the linearity (until very high magnetic fields) and reduces the thermal drift. All in all, the flexibility of graphene is a big plus and at the same time it makes a universal statement on the KPIs of graphene Hall elements difficult. Robust graphene Hall elements on SiO_2, e.g. with fixed high density through passivation, constitute good candidates to replace conventional magnetic field sensors in challenging environments (very low or high temperatures and wide dynamic field range with low linearity error). However, their performance does not promote their use against existing material systems in battery powered consumer devices requiring primarily a magnetic field resolution of few $\mathrm{nT}/\sqrt{\mathrm{Hz}}$. This requirement could be met, however, by large area, hBN-encapsulated devices if large scale fabrication of such heterostructure can be achieved in the future.

5.7.3. Comparison to other studies of graphene Hall element and other materials

The above given specifications allow us to conduct a comparison with other publications on graphene Hall elements as well as different Hall transducer materials such as Si, GaAs or InSb. For a fair comparison, however, only studies that investigate the performance of Hall transducers without any optimization through integrated

on-chip components and techniques, as described in Section 5.2, are considered[35]. This also impedes a direct comparison with data sheets of commercial Hall sensors that are typically integrated devices.

Other studies of graphene Hall elements

In the following comparison we focus on three key parameters at room temperature: The current related sensitivity (i.e. the Hall coefficient), the magnetic field resolution and the mobility which defines the resistivity for given density and reflects the quality of the material. As discussed in detail in the previous sections, the optimal density working point or range, n_{opt}, is subject to the fabrication method of graphene and its environment, i.e. the substrate and encapsulation material. This in particular the case for the magnetic field resolution. Moreover, for given charge carrier density, a large enough current has to be applied in order to reach the lowest possible value of the magnetic field resolution which corresponds to $B_{min,sat}$ in the 1/f noise regime. Unfortunately, the majority of studies in the literature does not discuss systematically whether the reported values are the best obtainable ones with respect to the accessible density and bias current range. In case of graphene Hall elements without gate, e.g. on SiC, as well as for other material systems, such as Si, GaAs or InSb discussed below, where the doping is fixed during fabrication of the material, the density dependence is not studied at all. Often, the density value for which the reported parameters were obtained is not specified explicitly. All in all, this makes a comparison between devices that are operated at different densities challenging. Therefore, we try to list as many reported results as possible in Table 5.3 and presume that these reported values correspond, in each case, to the best obtainable[36] or at least material typical[37] values. Moreover, we assume that the reported magnetic field resolution, $B_{min,meas}$, corresponds to the $B_{min,sat}$ value. This means that we presume that a large enough current was applied in all studies such that the devices were operated in the 1/f noise regime.

35 Moreover, we note that we consider here only Hall effect based magnetic field sensors. As discussed in Section 5.1, magnetic field sensors based on other physical effects, such as for instance TMR sensors, have to be compared on the system level taking into account many different aspects such as cost, power consumption, requirements by the specific application and more. This goes beyond the scope of this work.

36 In case of graphene devices equipped with a gate, we assume that the magnetic field resolution was measured over a larger density range and the best value was reported. However, most reports do not discuss explicitly an investigation of the density dependence.

37 In contrast to graphene, materials such as Si, GaAs or InSb cannot be fabricated with arbitrarily low doping while still remaining viably conductive. This limits the possibility for optimizing the sensitivity and magnetic field resolution.

The magnetic field resolution does not only depend on the applied current and the carrier density but also on the frequency (f_{meas}), the geometry (L/W ratio)[38] and the absolute dimensions of the measured device (i.e. the width W). These parameters vary substantially among the different studies. Therefore, for a proper comparison, all reported magnetic field resolution values will be normalized with respect to absolute dimensions and geometries and extrapolated to the same frequency using the Hooge formula for 1/f noise in case of negligible contact resistance contribution to the input resistance (see Eq. 5.20 in Section 5.3.2). The normalized magnetic field resolution value, $B_{min,norm}$, provided in the last column of Table 5.3 can then be calculated as follows

$$B_{min,norm}(f_{ref} = 1\,kHz) = B_{min,meas}(f_{meas}, n_{opt}) \cdot \sqrt{\frac{W}{L}} \cdot W \cdot \sqrt{\frac{f_{meas}}{f_{ref}}} \,. \quad (5.28)$$

We have chosen one kHz as reference frequency, f_{ref}, since most reported magnetic field resolution values were measured in the frequency range between one and a few kHz. n_{opt} is the density for which the magnetic field resolution was measured in the respective publication which we assume to be optimal in each report. No normalization is performed to account for different density working points since the functional dependency is complex and not all devices can be tuned over the same wide density range, as discussed above. In order for the Hooge model to be applicable for normalization and extrapolation, it is required that the reported magnetic field resolution was measured when the Hall element was operated unambiguously in the 1/f noise regime, such that $B_{min,meas}$ corresponds to $B_{min,sat}$ in the above equation. As shown in the previous sections, this is an accurate assumption for our few micrometer sized devices in the few kHz regime and over a wide density range. Moreover, almost all other studies on graphene Hall elements, which are listed in Table 5.3, provide data on the frequency dependence of the noise voltage, i.e. $N_V(f)$ curves. The data clearly show the expected frequency dependence of flicker noise in the few kHz range. Hence, all reported magnetic field resolution values in Table 5.3 correspond to the limit in the 1/f noise regime. As discussed in previous sections, the magnetic field resolution then no longer depends on the current when 1/f noise dominates. Hence, we do not have to account for the different bias currents

38 The current related sensitivity, i.e. the Hall coefficient, also depends on the geometrical prefactor G. However, it's value is between 0.7 and 1 for all used geometries (for squares and Hall crosses of different L/W ratios, respectively), as shown in Ref. 150. Hence, the Hall coefficient values in the comparison table are at most underestimated by 30% in case of square geometries compared to Hall cross geometries with $L/W \geq 3$.

Material	$S_{I,max}$ (V/AT)	μ (cm^2/Vs)	W (μm)	L/W	I (μA)	f_{meas} (kHz)	$B_{min,meas}$ (nT/$\sqrt{Hz}$)	$B_{min,norm}$ (nT μm /$\sqrt{Hz}$)
CVD-G/SiO_2 (Ref. 426)	310-1,200	n/a	5	6	3	30	43,000	480,755
CVD-G/SiO_2 (Ref. 242)	645-1,180	5,000	2	3	200	3	800	1,600
CVD-G/SiO_2 (Ref. 399)*	800	5,000	15	1.6	100	3	500	10,270
CVD-G/SiO_2 (Ref. 348)*	2,093	6,900	50	3.2	200	3	100	4,841
CVD-G/SiO_2 (Ref. 427)	1,800	n/a	1.5	3	3	1	20,000	17,321
CVD-G/SiO_2 (Ref. 430)	n/a	3,000	1	3	30	0.531	6,000	2,524
Epi-G/SiC (Ref. 431)	790	2,789	10	1	10	3.3	2,500	45,415
Epi-G/SiC (Ref. 398)*	87	3,000	150	3.3	3,000	1	95	7,805
Epi-G/SiC (Ref. 418)*	1,440	2,300	60	2	500	3	100	7,348
Exf-G/hBN (Ref. 349)	4,100-5,700	n/a	3	1	$\approx 10-100$	3	100	520
Exf-G/hBN (Ref. 397)	$\approx 5,000$	n/a	1	1	20	1	700	700
Exf-G/hBN this work	4,000-6,000	$\approx 70,000$	3	3	30	1	500	866
CVD-G/hBN (Ref. 198)*	1,986	1,600-3,700	120	2	10-1000	0.3	50	2,324

Table 5.3. Comparison of experimentally obtained parameters of graphene Hall elements on different substrates (all at room temperature). Different graphene fabrication methods, i.e. exfoliation, epitaxial growth on SiC as well as CVD growth with transfer on SiO_2/Si substrates, are distinguished which result in substantially different device dimensions (large-area devices are marked by *). We assume that the reported magnetic field resolution, $B_{min,meas}$, always corresponds to $B_{min,sat}$ in the 1/f regime such that the normalized magnetic field resolution, $B_{min,norm}$, can be determined using Eq. 5.28. See main text for further details and discussion.

that were used in the different experiments when comparing the reported magnetic field resolution values. Nevertheless, we include the bias current values in Table 5.3, if provided in the respective publication, because it demonstrates that larger bias currents are applied in case of devices with larger dimensions such that 1/f noise clearly prevails in the few kHz frequency range.

We emphasize that the definition of the normalized magnetic field resolution in Equation 5.28 is strictly valid only when the contact resistance is negligible (see

Section 5.3.2). As discussed in the two previous sections, this is not a good assumption for few micrometer sized, hBN-encapsulated devices. As a consequence, the obtained $B_{min,norm}$ value is overestimated (corresponding to an upper boundary) when compared to data obtained on large area devices or devices with smaller contact resistance contribution to the input resistance. In few cases in the literature, for instance by Schaefer and coworkers (Ref. 397), not only the measured B_{min} value is reported but also the width normalized magnetic field resolution, $B_{min} \cdot W$, at a certain reference frequency is provided for comparison. $B_{min} \cdot W$ corresponds to the above definition except for the additional $\sqrt{W/L}$ factor in the definition of $B_{min,norm}$ which takes different geometries into account.

In Sections 5.6.1 and 5.7.2, we found that $B_{min} \cdot W$ ranges between 500 and 1500 $nT\,\mu m/\sqrt{Hz}$ for hBN-encapsulated graphene Hall elements. For the Hall cross geometry with $L/W = 3$ as in our case, this results in $B_{min,norm}$ between 288 $nT\,\mu m/\sqrt{Hz}$ and 866 $nT\,\mu m/\sqrt{Hz}$. The upper (lower) boundary corresponds to the case of larger (smaller) contact resistivity. In Table 5.3, *this work* refers to typical values we experimentally obtained on and calculated for our hBN-encapsulated graphene Hall elements. We state the upper boundary of the $B_{min,norm}$ range as a conservative assessment. For graphene on SiO_2, we estimated $B_{min} \cdot W$ to range between several $1,000\,nT\,\mu m/\sqrt{Hz}$ and $20,000\,nT\,\mu m/\sqrt{Hz}$, which corresponds to $B_{min,norm}$ between $1,732\,nT\,\mu m/\sqrt{Hz}$ and $11,547\,nT\,\mu m/\sqrt{Hz}$ for $L/W = 3$ Hall cross geometry. In both cases, the values agree well with the results of the other studies listed in Table 5.3 when the respective reported magnetic field resolution as well as the respective design and operating parameters[39] are used to calculate $B_{min,norm}$.

For all devices on SiO_2 (CVD-G/SiO_2) and for hBN-encapsulated devices (Exf-G/hBN), the Si substrate was equipped with an integrated back gate to tune the density. Hence, the reported $S_{I,max}$ and $B_{min,meas}$ values can be considered as the best obtainable values over the density range which can be typically reached by means of electrostatic gating. Unfortunately, however, the majority of the studies does not provide a detailed discussion of the density dependence of both parameters. Moreover, for one and the same device, the ideal density working points with respect to highest sensitivity and best magnetic field resolution typically do not coincide, as discussed in the previous sections. In case of graphene on SiO_2 devices, the result of the first publication in Table 5.3 (Ref. 426) differs substantially from all other studies. Given the measurement at a frequency of 30 kHz using a current of only 3 μA, the measurement was likely not conducted in the 1/f noise regime. This means that

39 The geometry and dimensions were not explicitly stated in all studies. We then estimated the values from the provided micrograph images of the devices. In Ref. 349 the current value was not provided, so we estimated it from the given bias voltage range. However, in case of an operation in the 1/f regime, the exact current value is not relevant since B_{min} already approached $B_{min,sat}$.

the reported magnetic field resolution is much larger than the best obtainable value for larger current, i.e. $B_{min,sat}$. The most realistic conditions, comparable to today's commercial Hall elements, were considered in the studies on large-area devices (Refs. 348, 399, marked by an asterisk). The resulting $B_{min,norm}$ values are exactly within the range we predicted in our simulation of large-area graphene devices on SiO_2 using the key material parameters we obtained for exfoliated graphene devices on SiO_2. This shows that the quality of CVD grown and exfoliated graphene is similar on SiO_2 and, ultimately, determined by the substrate but not the fabrication process (when the transfer of CVD grown graphene is properly done and there are no major additional sources of contamination and disorder).

The $B_{min,norm}$ values for epitaxial graphene on SiC (second section of Table 5.3) are comparable to graphene on SiO_2. However, the density was fixed in all reports due to the absence of a back gate and was determined by the extrinsic doping of graphene on SiC. The large variation in the sensitivity values shows that the density working points, i.e. the level of extrinsic doping, were substantially different among the studied samples. Two publications investigated large-area devices for comparable operating parameters than in today's commercial Hall elements (Refs. 398, 418, marked by an asterisk). In agreement with our discussion in Section 5.7.2 and the results of our simulation of large-area devices with negligible contact resistance (panel d of Figure 5.14), the absolute difference in the obtainable $B_{min,sat}$ values is marginal when comparing intermediate to very high densities ($95\,nT/\sqrt{Hz}$ versus $100\,nT/\sqrt{Hz}$ for Ref. 398 and 418, respectively). However, as predicted by our simulation, the substantially lower sensitivity at very high density in Ref. 398 requires a six times larger current and also a larger device area than in Ref. 418 to reach the same $B_{min,sat}$ range. Finally, the $B_{min,norm}$ values are almost equal in both cases.

For hBN-encapsulated, exfoliated graphene Hall elements (third section of Table 5.3), the resulting $B_{min,norm}$ values are about one order of magnitude smaller than for graphene on SiO_2 or SiC. The typically obtained values in this work are similar to the findings of other studies (Refs. 349, 397). With respect to absolute values, the magnetic field resolution for these few micrometer sized devices is already comparable to large-area graphene on SiO_2. This shows the great potential of hBN-encapsulated graphene for which $B_{min,sat}$ in the $1-10\,nT/\sqrt{Hz}$ range is expected to be reached for large-area devices according to our calculations in Section 5.7.2. However, as will be discussed in the following Section 5.8, the fabrication of large-area, hBN-encapsulated graphene devices of similar quality than for exfoliated flakes yet cannot be realized. Mainly because of the low quality of CVD grown hBN films and the need for several transfer processes which leads to a large amount of residues. The authors of Ref. 198 reported the so far best results for CVD grown graphene on 20

nm thick, CVD grown hBN (i.e. without encapsulation[40]). Even though the lowest of all reported magnetic field resolutions in the literature is obtained (in absolute value), the normalized magnetic field resolution, $B_{min,norm}$, corresponds only to the lower boundary of the typical range for graphene on SiO_2, leaving room for further improvement.

Comparison to other Hall transducer materials

A fair comparison to other materials which are typically used as Hall transducer is even more difficult than comparing different graphene Hall elements among each other. In most cases, state of the art Hall sensors are integrated devices equipped with additional on-chip components that improve the performance. Only few studies provide data on Hall transducer level at room temperature and, then, mostly for micrometer-scale devices. Moreover, the reported values for one and the same material vary notably. Since the density is typically fixed by the doping of these semiconducting materials during fabrication, the density dependence cannot be investigated easily over a wide range. Finally, for the different studies on graphene Hall elements, the geometry, absolute dimensions and operating parameters (i.e. frequency and current) differ among the reports we considered for the following summary of specifications for bare Hall elements (based on Refs. 114, 397, 418 and references therein, as well as Refs. 150, 242, 379, 432–435). Typically, the magnetic field resolution values stated below refer to operation in the 1/f noise regime at a frequency of about one kHz.

Si Hall devices usually offer a current related sensitivity of around $100\,\Omega/\mathrm{T}$ [150, 242, 418] and have a quite low room temperature mobility of about $1,500\,\mathrm{cm}^2/\mathrm{Vs}$ [432]. Since the fabrication process of Si devices is very mature, the device quality is high and the 1/f noise level is low. Hence, magnetic field resolution values in the range between 30 and $100\,\mathrm{nT}/\sqrt{\mathrm{Hz}}$ are obtained [379, 432, 433]. Assuming a typical footprint of $100\,\mu\mathrm{m}$, it follows that $B_{min} \cdot W$ ranges between $3,000\,\mathrm{nT}\,\mu\mathrm{m}/\sqrt{\mathrm{Hz}}$ and $10,000\,\mathrm{nT}\,\mu\mathrm{m}/\sqrt{\mathrm{Hz}}$ (see also Ref. 397). Moreover, the great advantage is that integrated Si Hall sensors can be fabricated monolithically along with other on-chip components using standard Si CMOS processing steps.

40 Another study (Ref. 301) on all-CVD grown, hBN-encapsulated graphene Hall elements unfortunately did not specify the magnetic field resolution. However, due to high extrinsic doping and disorder that result from residues of the multiple transfer processes, only a low sensitivity is reported. This indicates a low quality of the device. Therefore, the publication was not included in the comparison table.

For Hall elements made from III-V compound semiconductors, such as GaAs, InAs, or InSb, the fabrication of integrated Hall sensors is more difficult and expensive (low compatibility to Si CMOS processes used for the other on-chip components). Moreover, in order to obtain high quality devices with good performance, these material systems are MBE grown which lacks high scalability and is expensive. As a consequence, III-V semiconductor Hall devices have succeeded rather in applications that require the best performance of Hall sensors but are not particularly cost sensitive. The room temperature mobility varies among the III-V semiconductors, from $\approx 7,000\,\mathrm{cm}^2/\mathrm{Vs}$ for GaAs to $\approx 50,000\,\mathrm{cm}^2/\mathrm{Vs}$ for InSb thin films [432, 434]. If such thin films are produced by cheaper large-scale fabrication methods, such as sputtering, instead of MBE growth, the mobilities are notably lower. In general, the obtainable current related sensitivies are higher than for Si and reach values of up to $1,100\,\Omega/\mathrm{T}$ in case of GaAs [433] and $2,750\,\Omega/\mathrm{T}$ in case of InSb [434]. $\mathrm{B}_{\mathrm{min}} \cdot \mathrm{W}$ at room temperature is approximately $1,000\,\mathrm{nT}\,\mu\mathrm{m}/\sqrt{\mathrm{Hz}}$ [397]. For dimensions on the order of $100\,\mu\mathrm{m}$, this corresponds to an absolute magnetic field resolution of about $10\,\mathrm{nT}/\sqrt{\mathrm{Hz}}$. For InSb, even around $1\,\mathrm{nT}/\sqrt{\mathrm{Hz}}$ could be reached theoretically [379, 432].

To conclude, graphene Hall elements on SiO_2 or SiC offer similar performance as Si with respect to magnetic field resolution, assuming an operation of graphene in the medium to high density regime. Even in this regime, however, the sensitivity of graphene Hall elements is still notably larger than for Si devices. As discussed above, our few micrometer sized, hBN-encapsulated devices are already comparable to large-area graphene on SiO_2 or SiC in terms of magnetic field resolution and, accordingly, also to conventional Si Hall devices with much larger footprint. Moreover, a similar magnetic field resolution as the best InSb devices can be reached given the same footprint. Maximum sensitivity and room temperature mobility are even superior. However, this requires that large-area, hBN-encapsulated graphene of high quality can eventually be fabricated in the future. Moreover, the fabrication of graphene has the potential to be cheaper than III-V materials. However, as for III-V materials, integrated graphene Hall sensors require an alignment between graphene growth and transfer and Si CMOS processes. The following and last section of this chapter discusses the current state of development of large scale fabrication as well as possible routes to integrate graphene with Si CMOS technology. In general, graphene offers the advantage that its density can be varied easily which offers a high flexibility to adapt to the requirements of any potential application of Hall sensors, e.g. low temperature drift or good linearity over a large magnetic field range when operated at high densities. At low density, record sensitivity values are obtained. Finally, apart from other materials for Hall sensors, graphene Hall elements also will have to compete with other types of magnetic field sensors that are based on different physical working principles, e.g. TMR sensors. Hence, graphene Hall

sensors have the potential to conquer a share in the Hall sensor market, if large scale fabrication and integration in Si CMOS are solved. However, they surely do not constitute a universal holy grail for magnetic field sensing taking both their current as well as potential future performance into account which were both discussed in this work. The performance of graphene Hall elements is promising but does not clearly outperform other existing solutions for magnetic field sensing, and in particular not in all aspects[41].

5.8. Realization of integrated, large-area graphene devices

In this last section, we review the status-quo of large-area fabrication of graphene and ideal substrates such as thin hBN films. Moreover, potential routes for and challenges in terms of Si CMOS integration are discussed. Finally, we assess the probability of occurrence that graphene becomes industrialized for mass market applications, including magnetic field sensing.

5.8.1. Large scale synthesis of high quality graphene

As discussed in Section 3.1, wafer scale graphene of reasonable electronic quality can be either epitaxially grown on SiC by selective thermal sublimation of Si or fabricated by chemical vapor deposition. For epitaxial graphene on SiC, the substrate is fixed and rather expensive while typically the initial doping of graphene is very high and the obtainable sensitivity values for Hall elements and mobilities are small [398]. In contrast, CVD growth of graphene allows for a flexible choice of substrate and, thus, potentially better transport properties and performance of Hall elements, depending on the substrate and the transfer processes. Direct growth of graphene on non-metallic substrates such as Si/SiO_2, GaN, Al_2O_3 or $SrTiO_3$, has been demonstrated using plasma enhanced CVD or MOCVD technology (for instance, 436–438). However, the intrinsic defect density is high and the graphene is polycrystalline. Electrical transport properties such as the mobility are worse than for graphene which is CVD grown on metallic foils and then transferred to

41 For instance, the potential cost per device with certain performance specifications even cannot be estimated here since the level of industrialization of graphene is still too early. Moreover, a detailed comparison on system level, i.e. comparing the final performance of integrated devices, with other magnetic field sensors such as TMR sensors, goes far beyond the scope of this work.

the same substrate material. Moreover, a high temperature is typically required for direct growth which makes the approach incompatible with Si CMOS pre-processed substrates (i.e. for back end of line integration).

CVD growth of graphene on catalytic substrates such as metal foils (mostly Cu or Pt [230, 246, 439]) or hydrogen-terminated Ge [232] results in graphene with the best electronic quality of all large-area fabrication processes, comparable to exfoliated graphene [203, 236]. The necessary transfer of graphene to the target substrate constitutes both an advantage and disadvantage. While such a transfer process may induce residues, cracks or wrinkles that lower the electrical properties and yield, it offers the possibility to choose an arbitrary substrate, including Si CMOS pre-processed wafers since it does not involve high temperatures. CVD growth of polycrystalline graphene and its transfer was demonstrated on meter scale [440–442] and single-crystalline graphene can be fabricated at least on cm scale and up to wafer scale [232–234, 247]. Moreover, there has been substantial progress in the development of transfer processes and handling of large-area graphene. Wet etching and polymer based transfer has become cleaner, bubbling transfer approaches allow reusing the catalytic substrate and, eventually, completely dry transfer methods were developed [245, 443, 444]. In particular, it was demonstrated that graphene can be picked up directly from the metal foils on which it was grown using exfoliated hBN flakes [203, 236]. Such devices then exhibit similar transport properties than exfoliated graphene on hBN. This shows that the fabrication of large-area graphene itself is no longer a substantial problem and all involved processes could be industrialized to increase the yield. The involved processes are more scalable and thus potentially cheaper than MBE growth of III-V compound semiconductors that compete in Hall effect based magnetic field sensing. However, it also shows that the physical properties of graphene are mainly governed by the target substrate. The best transport properties are obtained for hBN-encapsulated graphene, as demonstrated in the preceding chapters. Graphene Hall elements on SiO_2 or SiC do not realize the full potential of graphene and are not superior to existing Hall transducers materials, as discussed in the previous sections.

5.8.2. Large-area fabrication of ideal substrates and vdW heterostructures

The scalable fabrication of an ideal substrate and passivation layer is the key to establish large-area graphene successfully in commercial applications that benefit from the best possible electrical properties. As discussed in Section 3.1.3, an ideal

substrate or encapsulation to graphene is an insulating material which is atomically flat, chemically inert, has no dangling bonds and with absent remote-phonon scattering at room temperature. All these properties are unified in hBN. Up to few millimeter sized, bulk crystals are fabricated at very high temperatures and high pressure [282] or, as recently shown, also ambient pressure [445]. While these crystals are great for exfoliation of hBN flakes with randomly varying thickness, these crystals cannot be used directly as a substrate material due to their limited size and inhomogeneity. The ultimate goal is to find an approach that allows the growth of wafer-scale hBN thin films of about 10 to 20 nm thickness[42] either on a conventional Si/SiO_2 substrate or on metal foils which is then transferred to an arbitrary substrate.

The main approach to obtain such large-area hBN films is CVD growth (see reviews provided in Refs. 283, 446, 447). While monolayer hBN films can be grown on metal foils at large scale and mostly also with single-crystalline structure [288, 290, 295], the realization of large-area thin films of few atomic layers and up to 20 nm thickness remains challenging (see review focusing on multilayer hBN growth in Ref. 283). In particular, since the layer thickness should be homogeneous over the entire area. Unfortunately, the conditions for growth of multilayer hBN favor a polycrystalline structure as well as the formation of islands, steps and smaller hBN particles (white hBN dust). This increases the surface roughness and alters the dielectric properties. Moreover, based on our own experiments in CVD growth and transfer of hBN, even though the transfer methods developed for graphene are applicable to hBN films, the transfer results in more cracks and folds than for graphene. Also, the removal of polymer films used for the transfer results in more residues. Reasonable results of large-area multilayer hBN synthesis were achieved, for instance, on sapphire (Ref. 293), iron foils (Ref. 295) and directly on Si/SiO_2 wafers by means of plasma-enhanced CVD (Ref. 198). Comparing the absolute number of publications on CVD growth of graphene to synthesis of thin hBN films, reveals that the synthesis of hBN films was not yet been addressed to the same extent than the large-area fabrication of graphene. Hence, there is hope for future improvement.

The final step is then to fabricate large-area van der Waals heterostructures, e.g. in order to obtain large scale hBN-encapsulated graphene. This is typically realized by growing each layer separately and then stacking the layers on top of each other which, however, requires several transfer processes [448]. This may lead to a large

42 As discussed in Section 3.1.3, this is the minimum thickness for which graphene is screened properly from all extrinsic sources of disorder. Using such thin dielectric layers allows for better $n_g(V_g)$ stability and reproducibility. Moreover, in case of battery powered applications, the available maximum voltage is sufficient to tune graphene over a wide density range, offering more flexibility to set a suitable working point.

amount of residues, low yield due to inhomogeneity (cracks and folds) as well as degradation of the electrical properties of the final van der Waals heterostructure [301, 302]. The less transfer steps involved, the better the quality. For instance, for all CVD grown, large-area graphene on hBN, i.e. without encapsulation, the resulting device quality is already quite decent, as demonstrated for Hall elements in Ref. 198. Nonetheless, such devices are not yet comparable to heterostructures fabricated from exfoliated flakes using the dry pick-up method. Even though it was shown that CVD grown graphene can be picked up from the copper foil by an exfoliated hBN flake [203, 236], the concept could so far not be extended to larger area involving both CVD grown graphene and hBN. Recently, an approach using adhesive wafer bonding, as common in semiconductor manufacturing lines, was proposed [449]. It requires only the adhesive polymer for the first layer, which could be the top hBN layer, but then the other layers are successively stacked without being in direct contact to the adhesive polymer. However, it still requires that for each step the growth substrate is removed which is typically done by wet chemistry. As a consequence, the resulting transport properties for graphene/hBN are not yet convincing.

The ideal method to fabricate large-area van der Waals heterostructures would be to grow them directly with single-crystalline quality, layer by layer, without the need for any transfer process. This would result in atomically clean interfaces and, accordingly, best transport properties. Yang and coworkers (Ref. 450) showed that graphene can be grown directly on hBN flakes by plasma enhanced CVD. However, the mobility at low temperature was only about $5,000\,\mathrm{cm}^2/\mathrm{Vs}$. Another study (Ref. 451) reported that hBN and graphene can be grown simultaneously by a co-segregation method on iteratively boron, nitrogen and carbon doped Ni films. No transport properties were reported and the obtained hBN thickness was at most few atomic layers which is not enough for proper encapsulation. Finally, Wang and coauthors (Ref. 452) investigated the catalytic transparency of hBN on copper for CVD growth of graphene. They found that graphene grows nicely when the Cu foil is covered by only a very thin, i.e. single-atomic, layer of hBN. However, the growth rate drops fast and eventually graphene does not form at all when the hBN thickness is increased to few nanometer. Again, hBN is too thin to properly shield graphene from extrinsic disorder. To conclude, the optimal approach to grow van der Waals structures directly has not yet been found. Moreover, the large scale synthesis of hBN of sufficient thickness, to be used as a substrate, has not yet reached the same quality level than the fabrication of large-area graphene. In addition, the need for several transfer processes lowers the yield and electrical properties. To our knowledge, none of these processes has been industrialized so far, even though some approaches such as wafer bonding exhibit potential for doing so.

5.8.3. Si CMOS integration and other remaining challenges

It is clear that graphene or van der Waals heterostructures will not succeed without being combined with integrated circuits, i.e. merged with Si CMOS, since this offers a wide spectrum of additional functionalities and will favor a widespread use. For instance, all techniques that are described in Section 5.2, such as the spinning current method, and used to improve the performance of Hall elements require additional integrated circuits. Like for III-V semiconductors, the best approach is to integrate graphene or vdW heterostructures at the back end of line (BEOL) of Si CMOS fabrication, which means on top of the processed wafer. Since the transfer of large-area 2D materials does not involve high temperature steps, this is easily possible. Ref. 272 shows a realization of Si CMOS integrated graphene Hall devices, where graphene is transferred on top of the Si CMOS passivation layer, Si_3N_4, and then passivated by a polymer film (SU8). Similar concepts for Si CMOS integration are further discussed in Refs. 453–456. In all cases, the proper passivation or encapsulation of graphene is the key to guarantee a stable working point. Moreover, the dielectric environment has a significant impact on the electrical properties of graphene, i.e. hBN-encapsulation being the most favorable approach in this regard, as well as the 1/f noise level [408, 409]. Graphene can also be passivated by polymer films or oxides, but these degrade its performance [272, 457]. There is also the possibility to seal graphene with an air gap, i.e. burying it in the Si CMOS passivation layer [453]. Another issue which was raised by Lupina and coworkers (Ref. 458) is that the current and most promising way to fabricate large-area 2D materials, i.e. CVD growth on metal foils such as copper, results in a high level of residual metallic contaminants of the transferred graphene. While this plays a minor role when graphene is integrated with Si CMOS at the BEOL, it may lead to severe problems when front end of line (FEOL) approaches are considered.

All in all, the stage of industrialization is still early and there are no Si CMOS integrated graphene devices as a product on the market yet. For instance, the first commercially available graphene Hall sensors from Graphensic or Paragraf are analog devices, i.e. just the Hall element. At best, the chip carrier with the graphene element is attached to a PCB which offers additional functionalities (for demonstration purposes, no miniaturization yet). Given the high cost to develop integrated graphene devices, it is clear that a platform approach must be followed. The Hall sensor market alone is too small to drive such development, given also that the ultimate performance of graphene Hall devices relies on the large-scale fabrication of ideal substrates which still lags behind. In particular, electronic and optoelectronic applications will benefit from the excellent transport properties of high quality graphene, while other potential applications, e.g. battery electrodes, could be realized with

other large scale fabrication methods such as liquid-phase exfoliation for which the quality of graphene is lower. Use cases with similar requirements and complexity as integrated graphene Hall sensors include RF and THz applications [453, 454], broadband image sensors [459], MEMS applications [456], photonics [460], gas and bio sensors and others [221]. Moreover, the platform scenario should extend beyond graphene and consider also other 2D materials or a combination of them in van der Waals heterostructures in order to engineer new material properties [15]. Given the remaining challenges, the probability of occurrence of 2D materials in demanding applications is low on short term. To achieve these goals on a realistic time horizon of more than five years from now, the focus of large government funded research efforts, such as the Graphene flagship, must be adapted accordingly.

Looking specifically at the remaining challenges for graphene Hall devices, an industrialization of the processes, large-area fabrication and Si CMOS integration will help to reduce the sample to sample variation significantly. As discussed in the previous sections of this chapter, we observed a variation in the contact resistance among different samples as well as a sample-specific temperature dependence of the contact resistance. Moreover, the initial doping and the change of V_{CNP} with temperature are specific to each device. Finally, we found the density dependence of the offset to vary from sample to sample. For large-area devices, the contact resistance contribution to the input resistance becomes negligible. For clean processes, the initial doping is low and when 10-20 nm thick hBN is used as the only dielectric for field effect gating, and not also SiO_2, the thermal drift and the hysteresis are substantially reduced. Finally, for integrated devices, the offset can be almost completely eliminated using the spinning current technique. This also lowers the 1/f noise contribution in the frequency range below the switching frequency of the spinning current. The design of the integrated circuits can then also be adapted, for instance allowing for a higher switching frequency which accounts for the higher corner frequency of 1/f noise in graphene devices compared to Si devices. This would pave the way for high quality, hBN-encapsulated graphene to become a strong competitor to III-V semiconductor Hall devices with improved performance and with the additional advantages that graphene devices are cheaper to produce (CVD versus MBE) and the density working point can be easily tuned through electrostatic gating (for thin hBN dielectric, the available voltage in battery driven devices is sufficient). This offers exceptional flexibility to meet the requirements of different uses with a single type of Hall sensor.

Chapter 6

Summary

In this thesis, we investigated the reliable fabrication of high quality graphene and its use as Hall transducer material. Charged impurities and random strain fluctuations were identified as main culprits that deteriorate the electrical properties of graphene devices. It was shown that these extrinsic sources of disorder can be reduced through optimized device processing steps as well as the use of a proper substrate material for graphene such as hexagonal boron nitride (hBN). This insulating material is atomically flat and possesses a very low intrinsic density of charged impurities. By performing Raman spectroscopy and electrical transport measurements, both without and with applied magnetic field, on a large number of different types of graphene devices, it was demonstrated that the encapsulation of graphene between hexagonal boron nitride thin films is the best way to obtain high quality graphene devices. However, even for these hBN-encapsulated devices, we still observed a notable sample-to-sample variation of the electrical properties. Therefore, we developed a post-processing technique that allows us to improve the electrical properties of such devices both significantly and reliably. Since our technique is applied after device fabrication, we could also demonstrate its beneficial effect by comparing one and the same device before and after treatment.

We then assessed the application of such high quality graphene as Hall transducer material. The dependencies on and between all relevant operating parameters were explored. This allowed us to develop a deep understanding and empirical model for graphene Hall elements, including the interplay between thermal and 1/f noise in these devices. All key performance indicators for Hall sensors were measured and their typical values reported. For comparable device dimensions, hBN-encapsulated graphene Hall elements were found to have the potential to become a strong competitor to existing materials that are used in today's commercial Hall sensors. Unfortunately, the large-scale fabrication of hBN thin films still remains an unresolved challenge for the industrialization of large area, high quality graphene Hall elements.

Also, the Si CMOS integration demands further development. Even though the application of graphene in Hall devices is promising, as shown in this work, this use case alone does likely not justify the significant efforts and investments we expect to be necessary to industrialize the fabrication of high quality graphene devices. Instead, these efforts and costs must be shared by developing a common technology platform for 2D materials that can address several commercially attractive applications where graphene or another 2D material offers superior performance as well. We hope that the insights provided in this work can help to accelerate this process.

After this abstract, a more detailed summary of each chapter and its main findings is given in the following.

Chapter 2 gave a comprehensive review of the theory and experimental studies of electronic transport in graphene. We discussed the peculiar, linear band structure as well as many body interactions at low energy levels, near the CNP, that can lead to a renormalization of the Fermi velocity and energy dispersion. An overview of all relevant sources of extrinsic disorder was provided, with focus on charged impurities and random strain fluctuations as the main culprits for deteriorating the electrical properties of graphene. Both reduce the charge carrier mobility and cause the system to break down into electron hole puddles near the CNP. Due to the lack of a band gap, electrons and holes are also generated by thermal activation. We introduced the channel approximation model which requires only one ad-hoc parameter, the residual total charge carrier density at the CNP, to describe the crossover from the single to the two carrier type regime. This channel approximation was combined with the two carrier type Drude model to serve us as one of the main building blocks to describe graphene Hall elements.

In **Chapter 3**, different production, handling and device processing steps were described and compared. In particular, the assembly of van der Waals heterostructures, such as the encapsulation of graphene with thin hBN films, was discussed. We provided strong experimental evidence, based on Raman spectroscopy and electrical transport measurements on a large number of devices, that the properties of hBN-encapsulated graphene devices can be improved significantly and reliably by the post-processing technique which we proposed in this work. As first step, thermal annealing of the heterostructure at 500 °C directly after its assembly was found to be an important prerequisite for our method to have a significant effect. The second step is then performed after the device processing. We used an AFM tip in contact mode to mechanically treat the device. While the annealing facilitates the self-cleansing effect of the interfaces of the heterostructure, our hypothesis is that contact AFM irons out nanometer scale corrugations at the interfaces between hBN and graphene. After the treatment, hBN-encapsulated devices showed the smallest sample to sample variation in our Raman studies and exhibited the smallest width

of the 2D peak which is known to scale with the amplitude of random strain fluctuations. Finally, electrical transport measurements were found to be even more sensitive to the achieved improvement of our devices than Raman measurements. We demonstrated that we eventually reach the intrinsic limits of graphene when applying our post-processing technique. At room temperature, the obtainable mobility, limited by acoustic phonon scattering, was found to be about $80,000\,\mathrm{cm}^2/\mathrm{Vs}$ in the high density regime, e.g. at $10^{12}\,\mathrm{cm}^{-2}$. The residual total charge carrier density at the CNP, $n_{t,0}$, was limited by thermal activation of charge carriers and was typically around $7 \cdot 10^{10}\,\mathrm{cm}^{-2}$ at room temperature.

Chapter 4 discussed magneto-transport measurements in classically weak magnetic fields, i.e. when $\mu^2 B^2 < 1$ holds. The Hall coefficient was determined as a function of gate voltage, i.e. gate controlled average charge carrier density. We demonstrated that the two carrier type Drude model, in combination with the channel approximation, nicely describes the $R_H(V_g)$ dependence when the residual total charge carrier density is used as fitting parameter. At room temperature, maximum Hall coefficients of about $4500 - 5000\,\Omega/\mathrm{T}$ were typically obtained for hBN-encapsulated devices after applying the AFM ironing technique. For the lower bound, this corresponds to $n_{t,0} \approx 7 \cdot 10^{10}\,\mathrm{cm}^{-2}$. Again, strong statistical evidence was provided that AFM ironed devices exhibit superior performance to untreated, hBN-encapsulated devices or non-encapsulated devices on SiO_2. However, even for these monolayer graphene devices on SiO_2, subject to strong extrinsic disorder, we obtained larger Hall coefficient values than for hBN-encapsulated bilayer devices at room temperature. This is because of the different band structure and stronger thermal activation of charge carriers around the CNP in bilayer graphene. Hence, in terms of maximum sensitivity, monolayer graphene is clearly favorable.

According to the conventional tight binding model, which is based on a single particle picture, the linear energy dispersion in graphene can be characterized by a constant Fermi velocity of about $v_F \approx 1 \cdot 10^6\,\mathrm{m/s}$. Since the thermal activation of charge carriers scales with T^2/v_F^2, significantly larger $n_{t,0}$ and smaller $R_{H,max}$ values are expected at room temperature than we observed in our experiments. This suggests that the assumption of a trivial linear band structure is incorrect. In order to investigate this further, we measured the $R_H(V_g)$ dependence at different temperatures between room and liquid helium temperature. From these data we obtained the temperature dependence of the residual total charge carrier density, $n_{t,0}(T)$. Using the Fermi velocity as fitting parameter, but still as an energy or density independent parameter, we found a larger Fermi velocity value between 1.6 and $1.7 \cdot 10^6\,\mathrm{m/s}$ to describe the $n_{t,0}(T)$ data. We discussed that this could be explained by considering a renormalization of the band structure arising from many body interactions at low energy, i.e. in the vicinity of the CNP. This effect was found to be stronger for graphene devices with very low extrinsic disorder, as for instance in

our hBN-encapsulated devices.

Moreover, we performed temperature dependent electrical transport measurements in the absence of a magnetic field. We showed that, at high density, the increase in resistivity with temperature is mostly governed by acoustic phonon scattering. The maximum resistivity at the CNP was found to first decrease with temperature due to the thermal activation of additional charge carriers that become available for electrical transport. Eventually, this trend reverses around room temperature when optical phonon scattering prevails.

For the sake of completeness, the temperature dependent measurements were repeated for bilayer graphene devices. At room temperature, the magnetic field sensitivity did not exceed $1000\,\Omega/\mathrm{T}$ even for hBN-encapsulated devices. Again, we saw experimental evidence for an unexpected temperature dependence of $n_{t,0}$. We also acquired temperature dependent data on one monolayer graphene device whose crystallographic plane was closely aligned to that of the underlying hexagonal boron nitride. This causes a Moiré superlattice to emerge. This weak periodic potential induces replicas of the original charge neutrality point (ONP), the so called secondary neutrality points (SNP). We repeated the above analysis, using the Fermi velocity as variable parameter to fit the $n_{t,0}(T)$ data, and found that the thermal activation of charge carriers is much stronger at the SNP than at the ONP. We found v_F^{SNP} to be $0.6 \cdot 10^6\,\mathrm{m/s}$, in accordance with elsewhere reported scanning tunneling spectroscopy experiments.

Finally, the magneto-resistance at the CNP and its temperature dependence was studied as well. It exhibits a quadratic magnetic field dependence and arises from the coexistence of electrons and holes. We concluded that it is rather impractical to use this effect to measure magnetic fields. The quadratic magnetic field dependence implies a low signal level compared to the noise which reaches a maximum at the CNP. Since V_{CNP} was found to shift with temperature, it is challenging to keep the system stable at the CNP. In addition, the resistance at the CNP, that means the reference for the magneto-resistance, depends strongly on the temperature.

In **Chapter 5**, a comprehensive technical assessment of graphene based Hall elements was conducted. The key performance indicators of Hall sensors, such as sensitivity, magnetic field resolution, non-linearity and offset, were discussed and compared to other materials which are frequently used in Hall elements (e.g. Si, GaAs or InSb). We concluded that only high quality, hBN-encapsulated graphene has the potential to outperform most of these competitors, head-to-head with InSb in terms of performance but potentially cheaper to produce in high volume. For hBN-encapsulated devices, the channel width normalized magnetic field resolution, $B_{min} \cdot W$, was found to range between 500 and $1,500\,\mathrm{nT}\,\mu\mathrm{m}/\sqrt{\mathrm{Hz}}$ (depending on the contact resistance). However, in order to be able to fully leverage this promising performance in commercial Hall devices, the remaining challenges in the large-area

fabrication of thin hBN films and Si CMOS integration have to be resolved. The noise level is still too high and the magnetic field resolution too coarse in case of the hBN-encapsulated devices with only few micron dimensions which can be fabricated today from exfoliated flakes. The performance of non-encapsulated devices on SiO_2 was found to be less competitive in terms of obtainable magnetic field resolution ($B_{min} \cdot W$ ranged between several $1,000$ and $20,000\,nT\,\mu m/\sqrt{Hz}$). However, even these devices on SiO_2 still feature a significantly larger sensitivity than Si or GaAs based Hall sensors. We provided a specification sheet listing the typical values for all relevant performance indicators, both in case of hBN-encapsulated devices as well as non-encapsulated devices on SiO_2.

In order to determine the magnetic field resolution we had to measure accurately the electrical noise in graphene Hall elements, i.e. the noise voltage in transverse direction between the Hall voltage contacts. For this purpose, we described our measurement setup for which several measures, such as a home-built sample box with integrated amplifiers and carefully selected instruments, were applied to obtain a very low instrumental noise floor. This allowed us to determine accurately both the Hall coefficient as well as the noise voltage spectral density as a function of all relevant operating parameter, i.e. density, frequency and bias current. Such systematic investigation of the entire parameter space enabled us to study all dependencies on these parameters as well as the interplay between thermal and 1/f noise.

We demonstrated the applicability of the Hooge model by verifying experimentally all implied functional dependencies on bias current, frequency, density and area. In case of our few micrometer sized devices, 1/f noise was found to dominate in the low frequency regime and over the entire density range. The typical corner frequency was at least several kHz and up to several 10 kHz near the CNP. The noise voltage in transverse direction showed a Λ-shaped density dependence, i.e. it had a maximum at the CNP. From our data we extracted also the Hooge parameters, α_H, for a large number of devices. Its value for hBN-encapsulated devices, $\alpha_{H,hBN} \approx 10^{-5}$, was two orders of magnitude lower than for non-encapsulated devices on SiO_2, $\alpha_{H,SiO_2} \approx 10^{-3}$. Again, this shows the importance of reducing extrinsic disorder. The lowest values were obtained for hBN-encapsulated devices after applying our AFM ironing technique.

Moreover, we discussed the optimal working point of graphene Hall elements in terms of density. The highest sensitivity, i.e. the largest Hall coefficient, was found when the gate controlled average density equals $n_{t,0}$. However, this is not optimal for many other key performance indicators such as the magnetic field resolution, nonlinearity or thermal drift. We concluded to operate hBN-encapsulated graphene Hall elements in the density range between 2.5 and $5 \cdot 10^{11}\,cm^{-2}$ which equals about 4 to 5 times $n_{t,0}$ at room temperature.

We proposed an empirical model to estimate the sensitivity, magnetic field resolution, input resistance and power consumption for any type of graphene Hall element

in terms of footprint, geometry and device quality (i.e. type of substrate). It consists of three buildings blocks: The two carrier type Drude model, the Hooge model and the channel approximation. We showed that only four material parameters of graphene are required as input for the device simulation. These are the residual total charge carrier density at the CNP, the mobility, the Hooge parameter and the contact resistivity. Using the proposed model, we predict the best obtainable magnetic field resolution to be in the $1 - 10\,\mathrm{nT}/\sqrt{\mathrm{Hz}}$ range for hBN-encapsulated graphene Hall elements having a comparable footprint than the transducers in today's commercial Hall sensors. Moreover, the model was found to reproduce the functional dependencies of the above key performance indicators on all operating parameters, such as current, density or frequency.

Finally, we reviewed the status-quo of large-area fabrication of graphene and thin hBN films as ideal substrate as well as the other remaining challenges such as Si CMOS integration. The probability and the realistic time scale of graphene becoming industrialized for high volume applications such as Hall sensors was discussed. Even though, hBN-encapsulated graphene is a promising candidate for Hall elements, we concluded that further applications where graphene (or other 2D materials) are really superior to existing materials have to be identified first before the time and (financial) effort to industrialize graphene in high volume applications can be justified.

Appendix A

Zusammenfassung

In dieser Dissertation untersuchten wir die reproduzierbare und zuverlässige Herstellung von Graphen mit hoher Qualität, also mit bestmöglichen elektrischen Eigenschaften, sowie dessen Verwendung als Basismaterial für Hall-Sensoren. Geladene Störstellen und zufällig variierende Deformationen des Graphen-Gitters wurden als extrinsische Hauptursachen für die Verschlechterung der elektrischen Eigenschaften von Graphen identifiziert. Durch Optimierung der Prozessierungsschritte und durch Auswahl geeigneter Materialien als Substrat für Graphen können diese äußeren Störquellen deutlich verringert werden. Die Einkapselung, also die beidseitige Verpackung von Graphen mit zwei dünnen Schichten hexagonalen Bornitrids ist entscheidend dafür, Graphen Bauelemente von hoher Qualität zu erhalten. Dies wurde durch Raman Spektroskopie und elektronische Transportmessungen, sowohl mit als auch ohne Magnetfeld, für eine große Anzahl an Bauelementen bestätigt. Hexagonales Bornitrid (hBN) ist atomar flach und weist eine sehr geringe Dichte an geladenen Störstellen auf, wodurch es ein ideales Substrat für Graphen darstellt. Es zeigte sich aber, dass nicht alle dieser mit hBN verpackten Bauelemente immer zuverlässig gute elektrische Eigenschaften aufweisen. Wir entwickelten daher eine zweistufige Technik, mit welcher diese Eigenschaften reproduzierbar und signifikant verbessert werden können, und das selbst noch nach der Fertigstellung des Bauelements. Da der zweite, entscheidende Schritt unserer Methode nach allen Prozessierungsschritten angewendet wird, konnten wir einen direkten Vorher-Nachher-Vergleich durchführen, um die Wirksamkeit unserer Methode zu bestätigen.

Den zweiten Schwerpunkt dieser Arbeit bildete die Untersuchung des konkreten Anwendungsfalls von Graphen als Basismaterial für Hall-Sensoren. Die Abhängigkeiten von und zwischen allen relevanten Betriebsparametern solcher Sensoren wurden gemessen und ausführlich analysiert. Dadurch konnten wir ein gutes Verständnis und ein empirische Gesamtmodell für Graphen basierte Hall-Elemente entwickeln, welches auch das Zusammenspiel zwischen thermischem und 1/f-Rauschen berücksichtigt. Für alle wichtigen Kennzahlen von Hall-Sensoren wurden die typischen

Werte für verschiedene Arten von Graphen Bauelementen bestimmt. Nur für den Fall der Einkapselung von Graphen mit hBN, und der damit einhergehenden bestmöglichen elektrischen Eigenschaften, konnte das Potential von Graphen bestätigt werden, die heute in kommerziellen Hall-Sensoren verwendeten Materialien wie Si oder GaAs aufgrund besserer Leistungswerte zukünftig ersetzen zu können. Weitere Voraussetzung hierfür ist, dass die Fläche des mit hBN eingekapselten Graphen Hall-Elements ähnlich groß wie die heutiger Hall-Elemente ist, da das elektronische Rauschen mit zunehmender Fläche des Bauelements abnimmt.

Leider stellt die großflächige Herstellung von dünnen hBN Schichten noch immer eine Herausforderung für die Industrialisierung von Graphen basierten Bauelementen mit hoher Qualität dar. Auch die Integration von 2D-Materialien wie Graphen in Si CMOS Prozesse ist noch nicht ausgereift. Dies bedeutet, dass die Anwendung von Graphen in Hall-Sensoren zwar grundsätzlich vielversprechend ist, aber diese Anwendung alleine noch nicht den signifikaten Aufwand und die Kosten rechtfertigt, welche aus unserer Sicht nötig sein werden, um die Fertigungsprozesse von hochqualitativem Graphen, d.h. mit hBN Einkapselung, zur Industriereife zu bringen. Dieser Aufwand und die nötigen Investitionen müssen auf weitere kommerziell vielversprechende Anwendungen, für welche Graphen oder ein anderes zwei-dimensionales Material deutlich bessere Leistungswerte als konkurrierende Materialsysteme zeigt, verteilt werden. Ziel muss daher die Entwicklung einer universellen Technologieplattform für zwei-dimensionale Materialien sein. Wir hoffen, dass die in dieser Dissertation erarbeiteten Erkenntnisse dazu beitragen können, eine solche Entwicklung nicht nur zu rechtfertigen sondern auch zu forcieren.

Nach diesem kurzen Überblick über die Dissertation werden nun im Folgenden die wichtigsten Ergebnisse der einzelnen Kapitel detaillierter zusammengefasst.

Kapitel 2 bietet einen umfassenden Überblick über die Theorie des elektronischen Transports in Graphen sowie über die wichtigsten experimentellen Ergebnisse in der Literatur. Ein Schwerpunkt war dabei die besondere lineare Bandstruktur von Graphen. Nahe des energetischen Nullpunkts des Systems, um den sogenannten Dirac Punkt oder auch Ladungsneutralitätspunkt (engl. Charge Neutrality Point, daher im Folgenden abgekürzt als CNP), können Vielkörper-Wechselwirkungen zu einer Renormierung der Fermi-Geschwindigkeit und damit der linearen Energiedispersion führen. Außerdem gaben wir eine Übersicht über alle wichtigen Formen von extrinsischen Störquellen, wobei der Schwerpunkt auf geladene Störstellen und zufällig variierende Deformationen des Graphen Gitters (engl. random strain fluctuations), welche durch die Rauheit des Substrats induziert werden, gelegt wurde. Diese beiden Störquellen sind hauptursächlich für die Verschlechterung der elektrischen Eigenschaften von Graphen. Durch stärkere Streuung wird die Ladungsträgermobilität

verringert und am Ladungsneutralitätspunkt wird das System schließlich bei genauerer Betrachtung räumlich inhomogen. Es entstehen sogenannte Elektron-Loch-Pfützen, wobei die durchschnittliche, d.h. räumlich gemittelte, Ladungsträgerdichte weiterhin Null beträgt. Aufgrund des Fehlens einer Bandlücke werden Elektronen und Löcher nahe des Ladungsneutralitätspunkts auch durch thermische Aktivierung erzeugt. Wir zeigten, dass die Gesamtladungsträgerdichte durch eine vereinfachte Näherung des Leitungskanals (engl. channel approximation model) abgeschätzt werden kann. Hierzu wird lediglich ein Modellparameter, die restliche Gesamtladungsträgerdichte am Ladungsneutralitätspunkt, $n_{t,0}$, eingeführt. Die Gesamtladungsträgerdichte bezieht sich auf den Absolutwert der Dichte aller Elektronen und Löcher, d.h. deren Summe, im Gegensatz zur durchschnittlichen Ladungsträgerdichte, welche die Differenz der beiden Ladungsträgerdichten beschreibt und durch Anlegen einer Gate-Spannung variiert werden kann. Es konnte gezeigt werden, dass dieses Modell den Übergang des Systems zwischen dem Bereich mit einem vorherrschenden Ladungsträgertyp, d.h. Elektronen *oder* Löcher, und dem Bereich um den Ladungsneutralitätspunkt wenn beide Ladungsträger, d.h. Elektronen *und* Löcher, koexistieren (und das sprichwörtlich nebeneinander in Elektron-Loch-Pfützen), gut beschreibt. Diese Näherung für den Leitungskanal wurde dann mit dem für zwei Ladungsträgertypen angepasstem Drude Modell kombiniert. Dieses Konstrukt bildet dann den Hauptbestandteil unseres empirischen Gesamtmodells zur Beschreibung von Graphen Hall-Elementen, welches im Folgenden noch genauer beschrieben wird.

In **Kapitel 3** wurden verschiedene Herstellungsmethoden und Prozessierungsschritte für Graphen basierte Bauelemente beschrieben und verglichen. Im Fokus stand dabei die Herstellung von van der Waals Heterostrukturen, wobei verschiedene 2D Materialien übereinander gestapelt werden. Die Verkapselung von Graphen mit zwei dünnen hBN Schichten ist hierfür ein Beispiel. Mittels Raman Spektroskopie und Transportmessungen, die an einer Vielzahl an Proben durchgeführt wurden, konnten wir nachweisen, dass sich die elektrischen Eigenschaften von mit hBN verpacktem Graphen durch die in dieser Arbeit vorgestellten Technik, welche am Ende des Fertigungsprozesses des Bauelements angewendet wird, deutlich und zuverlässig verbessern lassen. Voraussetzung für die erfolgreiche Anwendung dieser zweistufigen Methode ist zunächst das Ausheizen der Heterostruktur bei 500 °C unmittelbar nachdem die einzelnen Schichten übereinander gestapelt wurden. Der zweite Schritt der Optimierungsmethode erfolgt dann erst nach der Fertigstellung des Bauelements, also nach den Lithographie- und den anderen nötigen Prozessschritten für das Ätzen der Struktur und die Kontaktierung. Hierzu verwenden wir die Spitze eines Rasterkraft-Mikroskops im Kontaktmodus, wobei eine vergleichsweise große Kraft

eingestellt wird, um die Heterostruktur mechanisch zu bearbeiten (in Analogie zu einem Bügeleisen). Wir entwickelten eine Hypothese und lieferten experimentelle Hinweise darauf, wie sich die Verbesserung der elektrischen Eigenschaften durch diese zweistufige Technik erklären lässt. Durch das Ausheizen der Heterostruktur unmittelbar nach dem Übereinanderstapeln wird der Selbstreinigungseffekt zwischen den Grenzflächen der Heterostruktur beschleunigt, wodurch die Grenzflächen größtenteils von Verunreinigungen befreit werden. Wenn diese dann nahezu frei von Kontamination sind, wirkt die von der Rasterkraft-Mikroskop Spitze ausgeübte Kraft wie ein Bügeleisen und Unebenheiten und damit zufällige Deformationen zwischen den Schichten werden reduziert. Mit diesen beiden Schritten adressieren wir somit vorwiegend die beiden vorherig beschriebenen Hauptursachen für eine Verschlechterung der elektrischen Eigenschaften von Graphen, geladene Störstellen und zufällige Deformationen. Mit hBN verkapselte Proben zeigten nach Anwendung unserer Methode die geringste Variation bei Raman Messungen zwischen verschiedenen Proben und die geringste Breite des 2D-Peaks, der bekanntermaßen mit der Amplitude der zufälligen Deformationen von Graphen skaliert. Mit elektronischen Transportmessungen konnten wir die erzielten Verbesserungen sogar noch besser quantifizieren als mittels Ramanspektroskopie. So konnten wir zeigen, dass die mit unserer Methode erreichten Werte letztlich den intrinsischen Grenzwerten von Graphen entsprechen. Im Bereich hoher Ladungsträgerdichten um $10^{12}\,\mathrm{cm}^{-2}$ betrug die typische Ladungsträgermobilität ca. $80.000\,\mathrm{cm}^2/\mathrm{Vs}$ bei Raumtemperatur. Dies entspricht in etwa dem bestmöglichen Wert im Fall von Streuung an akustischen Phononen. Weiter wird die Gesamtladungsträgerdichte am Ladungsneutralitätspunkt letztlich durch die thermische Aktivierung von Ladungsträgern begrenzt. Auch diesen Wert von etwa $7 \cdot 10^{10}\,\mathrm{cm}^{-2}$ bei Raumtemperatur erreichten wir durch Anwendung unserer Methode zuverlässig für alle Proben.

In **Kapitel 4** diskutierten wir die Ergebnisse unserer Magnetotransport-Messungen in schwachen Magnetfeldern bis etwa 100 mT, d.h. für den klassischen Fall wenn $\mu^2\mathrm{B}^2 < 1$ gilt. Der Hall-Koeffizient wurde als Funktion der Gate-Spannung, $\mathrm{R_H}(\mathrm{V_g})$, bzw. als Funktion der durchschnittlichen Ladungsträgerdichte, bestimmt. Wir konnten zeigen, dass das für zwei Ladungsträgertypen angepasste Drude-Modell in Kombination mit der Näherung für die Gesamtladungsträgerdichte im Leitungskanal die gemessenen $\mathrm{R_H}(\mathrm{V_g})$ Daten gut beschreibt, wobei nur ein Parameter, die restliche Gesamtladungsträgerdichte am Ladungsneutralitätspunkt, als Fit-Parameter verwendet wird. Bei Raumtemperatur betrugen die Maximalwerte des Hall-Koeffizienten etwa $4500 - 5000\,\Omega/\mathrm{T}$ für mit hBN verpacktes Graphen und nach Anwendung unserer oben beschriebenen Optimierungsmethode. Für den unteren Wert ergibt sich daraus $\mathrm{n_{t,0}} \approx 7 \cdot 10^{10}\,\mathrm{cm}^{-2}$. Durch Messung einer Vielzahl von Proben konnten wir erneut mit statistischer Aussagekraft zeigen, dass mit hBN eingekapselte Graphen Hall-Elemente nach Anwendung der Optimierungsmethode bessere Eigenschaften

aufweisen als ohne Einsatz dieser Technik, und dass generell mit hBN eingekapselte Bauelemente denen ohne Einkapselung auf einem SiO_2 Substrat überlegen sind. Doch selbst für einlagiges Graphen auf einem SiO_2 Substrat, d.h. ohne Einkapselung und damit einer Vielzahl an Störquellen ausgesetzt, konnten wir deutlich größere Hall-Koeffizienten als für zweilagiges Graphen, das mit hBN verpackt ist, messen. Dies erklärt sich durch die unterschiedliche Bandstruktur von zweilagigem Graphen und der damit einhergehenden stärkeren thermischen Aktivierung von Ladungsträgern nahe des CNP. Einlagiges Graphen ist daher hinsichtlich der maximal erzielbaren Magnetfeldempfindlichkeit eindeutig im Vorteil, weswegen wir uns in dieser Arbeit darauf fokussierten.

Gemäß dem konventionellen Tight-Binding Modell, das auf einem Einzelteilchenbild beruht, kann die lineare Energiedispersion in Graphen durch eine konstante Fermi-Geschwindigkeit von etwa $v_F \approx 1 \cdot 10^6\,m/s$ beschrieben werden. Da die thermische Aktivierung von Ladungsträgern dann mit T^2/v_F^2 skaliert, ergeben sich daraus bei Raumtemperatur deutlich größere Erwartungswerte für $n_{t,0}$ und kleinere für $R_{H,max}$ als wir aus der Analyse unserer Messdaten ableiten konnten. Dies deutet darauf hin, dass die Annahme einer vollständig linearen Bandstruktur nicht ganz richtig ist. Um dies weiter zu untersuchen, bestimmten wir die $R_H(V_g)$ Abhängigkeit bei verschiedenen Temperaturen zwischen Raumtemperatur und der Basistemperatur eines Flüssighelium-Kryostaten. Anhand dieser Daten wurde die Temperaturabhängigkeit der restlichen Gesamtladungsträgerdichte am CNP, $n_{t,0}(T)$, ermittelt. Bei Verwendung der Fermi-Geschwindigkeit als Fit-Parameter für diese $n_{t,0}(T)$ Daten, aber immer noch unter der vereinfachenden Annahme, dass die Fermi-Geschwindigkeit weiterhin unabhängig von der Energie bzw. Ladungsträgerdichte ist, ergab sich ein Wert zwischen $1,6$ und $1,7 \cdot 10^6\,m/s$ für die Fermi-Geschwindigkeit (bei einer linearen Dispersion). Anschließend wurde unter Berücksichtigung der Erkenntnisse aus anderen Experimenten in der Literatur diskutiert, inwiefern unsere Ergebnisse durch eine Renormierung der Bandstruktur erklärt werden können, welche durch Vielteilchen-Wechselwirkungen bei niedrigen Energien, d.h. in der Nähe des CNP, verursacht wird. Wir konnten zeigen, dass dieser Effekt bei Graphen Proben von hoher Qualität stärker ausgeprägt ist, also insbesondere im Fall unserer mit hBN eingekapselten Bauelemente nach Anwendung der zuvor beschriebenen Optimierungsmethode.

Anschließend führten wir temperaturabhängige Transportmessungen ohne Magnetfeld durch. Für unsere hoch-qualitativen Proben konnten wir zeigen, dass bei hoher Ladungsträgerdichte der kontinuierliche Anstieg des spezifischen Widerstands mit der Temperatur hauptsächlich durch Streuung an akustischen Phononen bestimmt wird. Der Widerstand am CNP nahm hingegen zunächst mit steigender Temperatur ab, was auf die thermische Aktivierung zusätzlicher Ladungsträger zurückzuführen ist, welche dann zur Leitfähigkeit beitragen. Ungefähr ab Raumtemperatur begann sich dieser Verlauf schließlich umzukehren, da dann die Streuung an

optischen Phononen zunehmend dominierend wird.

Der Vollständigkeit halber wurden außerdem die temperaturabhängigen Messungen für Proben mit zweilagigem Graphen wiederholt. Bei Raumtemperatur überstieg die Magnetfeldempfindlichkeit selbst im Fall von mit hBN gekapselten Proben nicht $1000\,\Omega/\mathrm{T}$. Auch hier lieferten die experimentellen Daten Hinweise auf eine unerwartete Temperaturabhängigkeit der restlichen Gesamtladungsträgerdichte am Ladungsneutralitätspunkt, $n_{t,0}(T)$. Außerdem untersuchten wir die Temperaturabhängigkeit von einer speziellen einlagigen Graphen Probe, wobei die Kristallgitter von der Graphen Schicht und von dem darunter liegenden hexagonalen Bornitrid absichtlich und mit nur einer sehr kleinen Winkelabweichung zueinander ausgerichtet waren. Dies führt zur Entstehung eines Moiré-Übergitters. Dieses schwache periodische Potential führt zu einer signifikanten Veränderungen der Bandstruktur. Neben dem ursprünglichen Dirac Punkt, dem originären Ladungsneutralitätspunkt (ONP), entstehen sekundäre Dirac bzw. Ladungsneutralitätspunkte (SNP). Wir wiederholten die obige Auswertung und passten unser Modell an die erhaltenen $n_{t,0}(T)$ Daten für beide Ladungsneutralitätspunkte separat an, indem wir jeweils unterschiedliche Fermi-Geschwindigkeiten als variable Parameter zuließen. Die thermische Aktivierung der Ladungsträger am SNP, d.h. die thermische Verbreiterung des Widerstand-Peaks am CNP, zeigte sich viel stärker ausgeprägt als am ONP. Unsere Auswertung lieferte $v_F^{SNP} \approx 0.6 \cdot 10^6\,\mathrm{m/s}$, in Übereinstimmung mit den Ergebnissen von Rastertunnelspektroskopie-Untersuchungen in der Literatur.

Am Ende des Kapitels wurde noch der magnetoresistive Effekt am CNP und dessen Temperaturabhängigkeit untersucht. Dieser entsteht durch die Koexistenz von Elektronen und Löcher am CNP. Der resultierende Magneto-Widerstand weist eine quadratische Magnetfeldabhängigkeit auf. Unserer Auffassung nach ist dieser Effekt allerdings weniger geeignet bzw. praktikabel, um Magnetfelder mit Hilfe von Graphen zu messen. Die quadratische Magnetfeldabhängigkeit führt zu einem schlechten Signal-Rausch-Verhältnis für kleinere Magnetfelder, da das Rauschen in Graphen am CNP maximal ist. Da sich die nötige Gate-Spannung, die angelegt werden muss um den Ladungsneutralitätspunkt zu erreichen, mit der Temperatur verschieben kann, ist es schwierig, den Arbeitspunkt stabil am CNP zu halten. Darüber hinaus ist auch der Widerstand am CNP, d.h. der Referenzwert für den Magneto-Widerstand, wie bereits diskutiert, stark temperaturabhängig.

In **Kapitel 5** wurde schließlich eine umfassende technische Bewertung von Graphen basierten Hall-Elementen durchgeführt. Die wichtigsten Leistungsindikatoren von Hall-Sensoren, wie Sensitivität, Magnetfeldauflösung, Nichtlinearität und Offset, wurden diskutiert und mit anderen Materialien verglichen, die bereits heute in Hall-Elementen verwendet werden (z.B. Si, GaAs oder InSb). Wir kamen zu dem Schluss, dass nur hoch-qualitatives, mit hBN eingekapseltes Graphen das Potential hat, hinsichtlich der wichtigsten Leistungsindikatoren die meisten dieser Materialien

zu übertreffen bzw. mit InSb gleichzuziehen, wobei sich Graphen aber voraussichtlich günstiger in großen Mengen produzieren lässt als InSb. Für mit hBN verpackte Bauelemente betrug die Magnetfeldauflösung nach Normierung mit der Breite des Hall-Elements, $B_{min} \cdot W$, zwischen 500 und 1.500 $nT\,\mu m/\sqrt{Hz}$ (abhängig vom Kontaktwiderstand). Um diese vielversprechende Leistungswerte in kommerziellen, großflächigen Hall-Elementen nutzen zu können, müssen jedoch die Herausforderungen bei der großflächigen Herstellung von dünnen hBN-Schichten sowie bei der Si-CMOS-Integration noch gelöst werden. Heute können Bauelemente aus mit hBN verpacktem Graphen nur mit Abmessungen von wenigen Mikrometern hergestellt werden, da bislang nur mit exfolierten hBN Flocken die nötige Qualität erreicht wird. Im Fall dieser kleinen Proben ist das elektronische Rauschen allerdings noch zu hoch und daher die Magnetfeldauflösung nicht gut genug. Die Leistungswerte von nicht-eingekapselten Graphen Bauelementen auf SiO_2 erwies sich als weniger konkurrenzfähig, zumindest hinsichtlich der Magnetfeldauflösung ($B_{min} \cdot W$ lag zwischen mehreren 1.000 und 20.000 $nT\,\mu m/\sqrt{Hz}$). Allerdings bieten selbst diese Bauelemente auf SiO_2 eine deutlich höhere Magnetfeldempfindlichkeit, also einen größeren Hall-Koeffizienten als Si- oder GaAs-basierte Hall-Sensoren. Unsere Ergebnisse über die wichtigsten Leistungsindikatoren von Hall-Sensoren wurden schließlich in einem vereinfachten Datenblatt für beide Arten von Graphen Hall-Elementen, d.h. für mit hBN eingekapselte Bauelemente und für Bauelemente auf SiO_2, zusammengefasst.

Um die Magnetfeldauflösung zu bestimmen, ist es nötig, das elektronische Rauschen zwischen den Hall-Kontakten, d.h. die Rauschspannung transversal zur Stromrichtung, genau zu messen. Zu diesem Zwecken stellten wir unseren Messaufbau vor, bei welchem durch verschiedene Maßnahmen ein sehr geringes Grund- bzw. Eigenrauschen des Messaufbaus erreicht werden konnte. Diese Maßnahmen umfassten unter anderem die sorgfältige Auswahl und Charakterisierung der Messinstrumente, die Implementierung einer sternförmigen Erdung aller Messinstrumente sowie die Verwendung einer selbstgebauten Messbox aus Metall, die mit einer Leiterplatte mit integrierten Verstärkerschaltungen ausgestattet ist. Auf diese Leiterplatte können die Proben gesteckt werden und die Hall-Spannung, also auch die Rauschspannung zwischen den Hall-Kontakten, wird vor der Übertragung an den Spektrumanalysator verstärkt. Damit konnten wir sowohl den Hall-Koeffizienten als auch die Rauschspannung in Abhängigkeit von allen relevanten Betriebsparametern von Hall-Elementen, also der Ladungsträgerdichte, der Frequenz und des Stroms, genau bestimmen. Diese systematische Untersuchung des gesamten Parameter-Raums ermöglichte es uns, alle Abhängigkeiten von und zwischen diesen Parametern zu ermitteln sowie das Zusammenspiel von thermischem und 1/f-Rauschen im Detail zu studieren.

Wir zeigten die Anwendbarkeit des Hooge-Modells zur Beschreibung des 1/f-Rauschens in Graphen Hall-Elementen, indem wir alle vom Modell vorhergesagten

funktionalen Abhängigkeiten von Strom, Frequenz, Ladungsträderdichte und Fläche experimentell bestätigten. Im Fall unserer nur wenige Mikrometer großen Bauelemente war das 1/f-Rauschen im Bereich niedrigerer Frequenzen über den gesamten untersuchten Ladungsträgerdichte-Bereich vorherrschend. Die typische Frequenz, ab welcher letztlich das thermische Grundrauschen erreicht wird, lag bei mindestens mehreren kHz und sogar bis zu mehreren 10 kHz in der Nähe des CNP. Die Rauschspannung zwischen den Hall-Kontakten zeigte eine Λ-förmige Abhängigkeit von der Ladungsträgerdichte, d.h. sie hatte ein Maximum am CNP. Aus unseren Rohdaten konnten wir außerdem den Hooge-Parameter, α_{H}, für eine größere Anzahl an Proben bestimmen. Der Hooge-Parameter war für mit hBN verpackte Graphen Proben zwei Größenordnungen kleiner, $\alpha_{\mathrm{H,hBN}} \approx 10^{-5}$, als für Graphen Proben auf SiO_2, $\alpha_{\mathrm{H,SiO_2}} \approx 10^{-3}$. Dies zeigt, dass es auch für das elektronische Rauschen wichtig ist, extrinsische Störquellen weitestgehend zu reduzieren. Entsprechend wurden für mit hBN eingekapselte Graphen Hall-Elemente nach Anwendung unserer Optimierungsmethode auch die niedrigsten Werte für den Hooge Parameter erzielt.

Außerdem diskutierten wir den optimalen Arbeitspunkt von Graphen Hall-Elementen in Bezug auf die Ladungsträgerdichte. Die höchste Magnetfeld-Sensitivität, d.h. der größte Hall-Koeffizient, wird erreicht, wenn die durchschnittliche Ladungsträgerdichte, welche von der Gate-Spannung bestimmt wird, den gleichen Wert wie die restliche Gesamtladungsträgerdichte am Neutralitätspunkt hat, d.h. wenn $\mathrm{n_g} = \mathrm{n_{t,0}}$. Für andere wichtige Leistungsindikatoren wie die Magnetfeldauflösung, die Nichtlinearität oder die Temperaturkoeffizienten ist dieser Arbeitspunkt jedoch nicht optimal. Wir kamen daher zu dem Schluss, dass mit hBN verpackte Graphen Hall-Elemente im Bereich einer durchschnittlichen Ladungsträgerdichte von $2,5$ bis $5 \cdot 10^{11}\,\mathrm{cm}^{-2}$ betrieben werden sollten, was etwa dem 4- bis 5-fachen des Werts von $\mathrm{n_{t,0}}$ bei Raumtemperatur entspricht.

Wir erarbeiteten auch ein empirisches Modell, um die Sensitivität, die Magnetfeldauflösung, den Widerstand des Bauelements und damit den Stromverbrauch für jegliche Art von Graphen Hall-Elementen abschätzen zu können, d.h. für eine beliebige Grundfläche und Geometrie sowie Qualität des Graphen Bauelements, welche primär durch das Substratsmaterial bzw. die Einkapselung bestimmt wird. Dieses Modell besteht aus drei Bausteinen: Dem für zwei Ladungsträgertypen angepassten Drude-Modell, der Näherung für die Gesamtladungsträgerdichte im Leitungskanal und dem Hooge-Modell. Wir konnten zeigen, dass vier Materialparameter von Graphen für die Simulation des Bauelements ausreichend sind. Diese sind die restliche Gesamtladungsträgerdichte am CNP, die Ladungsträgermobilität, der Hooge-Parameter und der (spezifische) Kontaktwiderstand. Mit Hilfe dieses Modells wurden dann unter anderem hBN-verpackte Graphen Hall-Elemente simuliert. Vergleiche mit den Messergebnissen bestätigten, dass das Modell die funktionalen

Abhängigkeiten der oben genannten Leistungsindikatoren von allen Betriebsparametern, wie Strom, Dichte oder Frequenz, reproduzieren kann. Im Fall einer vergleichbaren Grundfläche des mit hBN verpackten Graphen Hall-Elements wie in heutigen kommerziellen Hall-Sensoren, erwarten wir eine Magnetfeldauflösung im Bereich von $1 - 10\,\mathrm{nT}/\sqrt{\mathrm{Hz}}$, was einen sehr vielversprechenden Wert darstellt.

Schließlich gaben wir eine Übersicht über den Stand bei der großflächigen Herstellung von Graphen und dünnen hBN-Schichten sowie über die verbleibenden Herausforderungen bei der Integration von Graphen und anderen 2D Materialien in die Prozesse der Si-CMOS Fertigungstechnologie. Wir diskutierten die Eintrittswahrscheinlichkeit und einen realistischen Zeitrahmen für die Industrialisierung von Graphen und damit für Anwendungen mit hohen Stückzahlen wie z.B. als Hall-Sensoren. Obwohl mit hBN eingekapseltes Graphen ein vielversprechendes Materialsystem für Hall-Elemente darstellt, kamen wir zu dem Schluss, dass zunächst weitere Anwendungsmöglichkeiten identifiziert werden müssen, bei denen Graphen oder andere 2D-Materialien den bisherigen Materialien ebenfalls überlegen sind, bevor der zeitliche und finanzielle Aufwand für die Industrialisierung von Graphen bzw. einer universellen Plattform für die Großserienfertigung von 2D-Materialien und van der Waals Heterostrukturen gerechtfertigt werden kann.

Appendix B

Supplementary transport theory

B.1. Quantum effects in transport beyond the diffusive picture

Quantum interference effects

The wave-like nature of charge carriers becomes relevant when the phase information is maintained over a length scale L_ϕ significantly larger than the mean free path for elastic momentum relaxation, L_{mfp}. In the following discussion we will also assume that L_{mfp} is much smaller than the sample dimensions ($L_{mfp} \ll L, W$) to avoid the relevance of boundary scattering. Carrier trajectories may form a closed loop and, when time reversal symmetry is conserved in the absence of a magnetic field, two counterpropagating paths acquire a well-defined phase difference leading to quantum interference corrections to the Drude-Boltzmann conductivity (see also Section 2.1.5 and Figure 2.2d). In conventional 2DEG without strong spin orbit coupling, these paths interfere constructively enhancing the probability for backscattering (referred to as weak localization). As a result the resistivity (conductivity) is increased (decreased). When a weak magnetic field is applied, the time reversal symmetry breaks and the enhanced backscattering is suppressed. This results in a negative magneto-resistance (MR) or equivalently positive magneto-conductance (MC), which constitutes an experimental indication of weak localization. The amplitude and shape of the correction to the MC curve is governed solely by the phase breaking inelastic scattering time τ_ϕ and by the elastic momentum relaxation time τ (the transport scattering time in the Drude-Boltzmann formalism when the relaxation time approximation applies).

In graphene, however, this quantum interference effect due to time reversal symmetry is more rich and not solely determined by these two time scales, because the time reversal operation connects electronic states of the two different valleys and charge carriers are chiral in nature (as discussed in Section 2.1.5). The amplitude, shape and even whether weak localization (WL) or weak anti-localization (WAL) occurs also depends on the intervalley and symmetry breaking intravalley scattering times (τ_{iv} and τ_*, respectively). The correction to the magneto-conductance contains three terms and reads as follows [42, 461]:

$$\Delta\sigma(\mathrm{B}) = \frac{\mathrm{e}^2}{\pi\mathrm{h}}\left[\mathrm{F}\left(\frac{\tau_{\mathrm{B}}^{-1}}{\tau_\phi^{-1}}\right) - \mathrm{F}\left(\frac{\tau_{\mathrm{B}}^{-1}}{\tau_\phi^{-1}+2\tau_{\mathrm{iv}}^{-1}}\right) - 2\mathrm{F}\left(\frac{\tau_{\mathrm{B}}^{-1}}{\tau_\phi^{-1}+\tau_{\mathrm{iv}}^{-1}++\tau_*^{-1}}\right)\right] \quad \text{(B.1)}$$

with $\tau_{\mathrm{B}}^{-1} = 4\mathrm{eDB}/\hbar$ and $\mathrm{F}(\mathrm{z}) = \ln(\mathrm{z}) + \Psi(0.5+\mathrm{z}^{-1})$, where Ψ is the digamma function and $\mathrm{D} = \mathrm{v}_{\mathrm{F}}^2\,\tau/2 = \mathrm{v}_{\mathrm{F}}\,\mathrm{L}_{\mathrm{mfp}}/2$ the diffusion coefficient. Note, the formula is only applicable when the momentum relaxation time τ is the smallest in the system, in agreement with the assumed diffusive regime ($\mathrm{L}_{\mathrm{mfp}} \ll \mathrm{L}, \mathrm{W}$). Magneto-conductance measurements are performed at lower temperature and at higher density to avoid the very strong classical magneto-resistance due to the coexistence of electrons and holes at the CNP (see Section 2.3.3). Then, $\mathrm{k_B T} \ll \mathrm{E_F}$ holds and τ can be derived directly from the measured conductivity in the absence of a magnetic field (given by the Boltzmann model in Equation 2.48) [18, 77, 462]. Here, the density is well known from Hall measurements or from field effect measurements. Note that the correction to the conductivity value by this quantum interference effect is small ($\Delta\sigma_{\mathrm{QI}} \ll \sigma(\mathrm{B}=0)$), on the order of one percent, causing no significant error in the calculation.

In the remainder, we discuss each scattering rate in the above formula (Eq. B.1) in more detail. The phase breaking rate due to the magnetic field τ_{B}^{-1} scales with $\tau\mathrm{B}$. Together with the phase breaking rate by inelastic scattering τ_ϕ^{-1}, it determines the amplitude and width of the MC curve. In graphene, inelastic electron electron scattering was found to determine the phase breaking rate τ_ϕ^{-1} over a wide temperature range due to weak electron-phonon scattering [77]: It increases linearly (or even super-linearly) with temperature and also increases when the density is tuned to lower values (because electron electron interactions augment). It is also affected by the momentum scattering rate $\tau_\phi^{-1} \propto \tau^{-1}$ [77]. Taking all this information together, these dependencies imply that for cleaner graphene devices with higher mobility this quantum interference effect occurs for a much narrower range of magnetic fields of a few mT only [18]. Since τ_ϕ^{-1} determines the amplitude of the effect, the latter is more difficult to measure at higher temperatures and densities (which limits the window within τ_ϕ^{-1} can be tuned). Intervalley scattering, characterized by the rate τ_{iv}^{-1}, is accompanied by a large change in momentum and

is therefore caused exclusively by short range disorder. Because the time reversal operator connects the electronic states of both valleys, the carriers from different valleys can (always) interfere so that intervalley scattering restores weak localization [41, 462]. Scattering mechanisms that do not break time reversal symmetry are unable to cause direct backscattering ($\mathbf{k} \rightarrow -\mathbf{k}$) within a valley due to pseudospin protection and pseudospin-momentum locking. The intravalley symmetry breaking rate τ_*^{-1} encompasses all scattering mechanisms that lift this pseudospin protection which suppresses quantum interference *within one* valley [42, 77, 461]. Trigonal warping has been identified as one possible channel for intravalley symmetry breaking [463]. When higher order momentum terms are considered in the approximation of the band structure near K, the Dirac cone no longer exhibits perfect isotropy (at higher energies). This gives rise to a small probability for direct intravalley backscattering even for symmetry conserving long range disorder such as charged impurities [18, 25, 42]. This would imply that τ_*^{-1} becomes larger at higher carrier densities. Another, more likely explanation is given by long range disorder such as random strain fluctuations that breaks, inherently, effective time reversal symmetry within one valley as it is accompanied by a random pseudomagnetic field [18, 461, 462].

When both intervalley and symmetry breaking intravalley scattering are negligible ($\tau_*, \tau_{\mathrm{iv}} \rightarrow \infty$), the first two terms in the magneto-conductance formula cancel each other and the third (negative) term dominates resulting in weak antilocalization. This corresponds to the case where symmetry conserving (scalar) long range disorder dominates as discussed in Section 2.1.5. Experiments on graphene devices with low mobility supported by SiO_2 [462] as well as mediocre mobility devices fabricated on top of hBN (but not encapsulated by hBN so that they are exposed to polymers during processing) [18] reveal that the symmetry breaking intravalley scattering time τ_* is density-independent (which implies that trigonal warping is not the cause). It is also the shortest time scale approaching the momentum relaxation time $\tau_* \approx \tau$ (see fits to the data using Eq. B.1 in Figure B.1a,b). The intervalley scattering time on the other hand fulfills $\tau_{\mathrm{iv}} \gg \tau$. Hence, we conclude that short range disorder is weak. For measurements at liquid He temperatures, the phase breaking time is the largest in the system $\tau_\phi > \tau_{\mathrm{iv}}$. Then, WAL (interference within one valley) is suppressed and WL is restored because backscattering is enhanced by the intervalley scattering mediated constructive interference of time reversed trajectories in the two valleys. Tikhonenko et al. (Ref. 462) excluded trigonal warping and further random electrostatic potentials coming from randomly distributed charged impurities (a third explanation for asymmetry in the crystal structure) as the main source for intravalley scattering since their strength cannot account for $\tau_* \approx \tau$. Couto et al. (Ref. 18) followed this argumentation and concluded that a proper theoretical model of random strain fluctuations explains $\tau_* \approx \tau$. The assertion that random strain

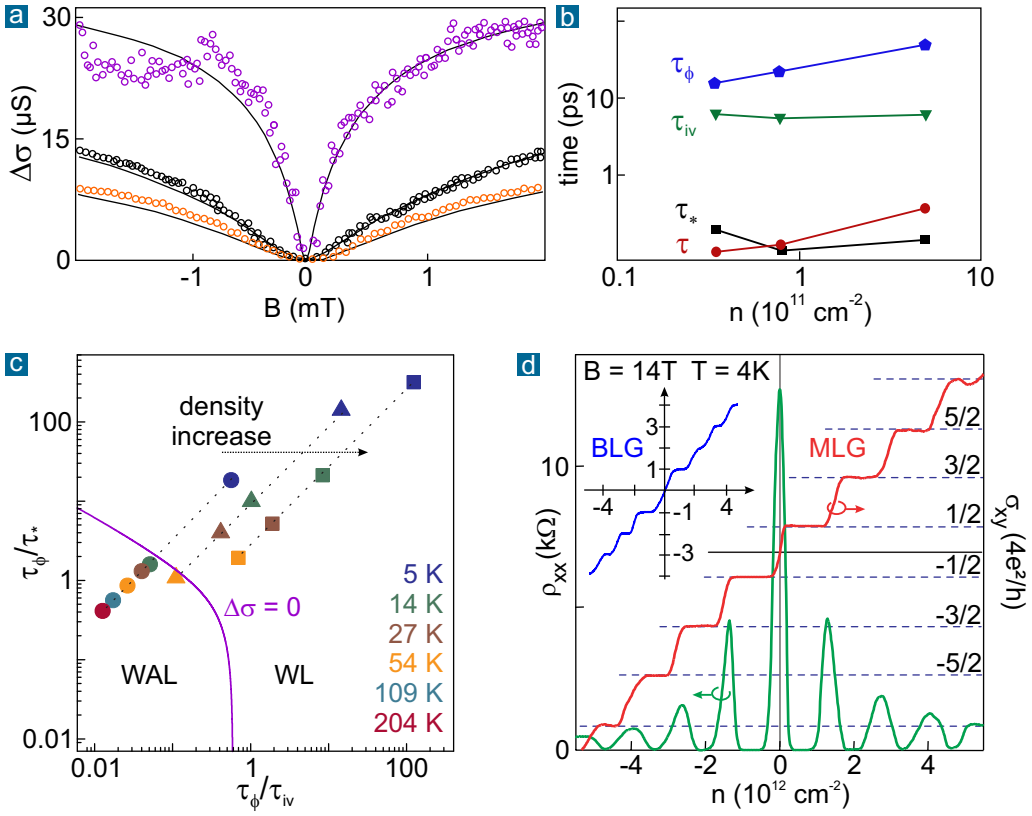

Figure B.1. Quantum interference and quantum Hall effect in graphene. a: Magnetoconductance data and fit curves based on Eq. B.1 for various carrier densities (T=250mK, graphene on hBN), note the small B field values. b: Density dependence of the elastic and inelastic scattering times used as fit parameters for the data in panel a. τ is derived separately from the Boltzmann conductivity. Panels a and b have been replotted from Ref. 18. c: Crossover between WAL and WL regimes in graphene on SiO_2, showing the tunability of τ_ϕ (modified from Ref. 77). d: Half-integer quantum Hall effect in monolayer graphene (MLG) and integer QHE with a step across zero density for bilayer graphene (BLG, inset), taken from Ref. 33.

fluctuations are the main culprit is also supported by several other experimental findings requiring long range disorder (see Section 2.3.2).

It was shown in another contribution by Tikhonenko et al. (Ref. 77) that weak antilocalization can occur even when $\tau_* \approx \tau$. By increasing temperature (τ_ϕ decreases) and decreasing density (τ_ϕ decreases, $\tau_{\rm iv}$ slightly increases) they were able to map the crossover between WL and WAL in graphene as it depends on the two ratios

τ_ϕ/τ_* and $\tau_\phi/\tau_{\mathrm{iv}}$ (see Figure B.1c). We can understand this recovery of weak antilocalization also from the following handwaving argument: The dephasing length becomes smaller at higher temperature and lower density. This length scale determines the number of possible loops that may give rise to constructive or destructive interference effects. Since sharp defects causing intervalley scattering are seldom and since fewer strain fluctuations are relevant for smaller loops, the ideal graphene case is restored and chirality constraints regain relevance. The number of loops, however, also determines the strength of the interference effect making it challenging[1] to detect WAL and WL under these experimental conditions, especially when mixed with other effects causing a change of the conductivity with magnetic field, such as for instance all classical effects (not only at the CNP), universal conductance fluctuations (UCFs), electron-electron interactions [76, 464], and the onset of Shubnikov de Haas oscillations. The latter already appear at very small magnetic fields for high mobility samples at low temperature.

Quantum Hall effects

The charge carriers in a two-dimensional system execute cyclotron motion, which in a strong perpendicular magnetic field ($\mu B \gg 1$) form closed orbits. The magnetic flux encompassed by the orbit produces a phase shift of the electron's wavefunction which must be an integer multiple of 2π. Hence, only discrete orbits are allowed and the energy spectrum is discretized into Landau levels each possessing a degeneracy equal to the number of elementary flux quanta penetrating the sample. This discrete Landau level spectrum manifests itself at lower magnetic fields in Shubnikov de Haas oscillations in the longitudinal conductivity and at higher fields in a quantization of the Hall conductivity. Plateaus form at integer values of the conductance quantum e^2/h, whenever a Landau level is completely filled, while the longitudinal conductivity becomes zero (integer quantum Hall effect) [7, 151]. In conventional parabolic 2DEG, the Landau levels are equidistant in energy. The linear energy dispersion in

1 As a consequence, all so far reported experiments relied on an averaging over a certain density range to suppress UCF with high magnetic field periodicity. Considering that recent hBN-encapsulated graphene devices show even higher mobility than the uncovered samples on hBN used by Couto et al. (Ref. 18), the magnetic field range in which these interference effects emerge drops even further to about or below one mT. This makes it almost impossible to record them in conventional cryostats with superconducting magnets which have an offset and hysteresis larger than this range (requiring at least a reference magnetic field sensor). At low temperature, those high mobility samples may enter the quasi-ballistic regime. Then, the condition $L_{\mathrm{mfp}} \ll L, W$ is no longer met and boundary scattering dominates the magnetoresistance, maybe even together with hydrodynamic effects predicted by theory [162, 209]. For these reasons, quantum interference measurements are not the scope of this thesis.

graphene and the chiral nature of its charge carriers (pseudospin, see Section 2.1.4) lead to a non-equidistant and peculiar Landau level spectrum [33]:

$$E_N = \mathrm{sign}(N)\, v_F \sqrt{2e\hbar B\,(N + 1/2 \pm 1/2)}\,, \tag{B.2}$$

where $\pm$ refers to the pseudospin describing the two sublattices. As a consequence of this Landau level spectrum, Shubnikov de Haas oscillations exhibit an 'odd' phase shift described by [27, 33]

$$\Delta R_{xx} = R(B,T)\cos(2\pi(B_F/B + 1/2 + \beta)). \tag{B.3}$$

The additional phase contribution β equals $1/2$ for a Berry phase of π (and zero for 2π (0), not relevant for monolayer graphene). The Berry phase affects the position of the plateaus in the Hall conductivity. They are located at *half-integer* filling factors $\nu = N + 1/2$. The Landau levels are fourfold degenerated due to the spin and valley degeneracy (see Figure B.1d) [27, 33]. A third type of the integer quantum Hall effect is realized in bilayer graphene which has a parabolic energy spectrum but chiral charge carrier with a Berry phase of 2π ($\beta = 0$). The latter is not reflected in an unusual sequencing of the plateaus at high density but leads to an eightfold degeneracy and a step of 8 conductance quanta in the Hall conductance as one crosses zero density [7, 465].

The quantum Hall effect in graphene is very robust because the spacing between the levels with low orbital index is large. It can even survive up to room temperature [466]. The energy gap between the lowest and first excited Landau level exceeds the thermal energy by a factor of ten at room temperature when a magnetic field of 30T is applied. In combination with its extraordinary high room temperature mobility, this makes graphene a very promising candidate for a non-cryogenic metrological standard. The steady improvements in the sample quality, which we discuss in Chapter 3 were mainly accomplished by the hunt for fractional quantum Hall ground states. The fractional quantum Hall effect was first demonstrated in two-terminal conductivity measurements on suspended devices [72, 73] and later confirmed by Hall plateau measurements on hBN-supported graphene [74]. The development of a new assembly technique for van der Waals heterostructures [10], which is described in detail in Section 3.1.4, allowed further improvements in terms of the hBN-graphene interface cleanness and the (tunable) contact resistance. Moreover, the implementation of thin graphitic gates offered screening of charged impurities and a more homogeneous spatial electric field for the field effect gating (see for instance Refs. 43, 47, 206, 326, 327 and Sections 2.2.5 and 3.1.4). All these experimental achievements enabled the confirmation of Hofstadter's butterfly spectrum with fractal quantum Hall [45, 50, 52] and even fractional fractal quantum Hall states [47] in Moiré superlattices. Such superlattices form naturally at clean interfaces between graphene and hexagonal boron nitride (see Section 2.1.6). When the

length scale of the superlattice becomes similar to the magnetic length, the commensurability condition for Hofstadter's self-similar energy spectrum can be easily fulfilled [45, 50, 52]).

B.2. Hall coefficient and resistivity in the two carrier type Drude model

We calculate the Hall coefficient and the longitudinal resistivity for coexisting electrons and holes in a two-dimensional system. The system is considered to be homogeneous and diffusive which allows to employ the Drude-Boltzmann theory (simple two carrier type model). The configuration as depicted in panel b of Figure B.2 is assumed: The magnetic field B has only one component in z-direction perpendicular to the two-dimensional system which is located in the xy-plane. The movement of the charge carriers can then be described by the following classical equations of motion, where e is the elementary charge and taken to be positive:

$$\mathrm{m_e}\dot{\mathbf{v}}_\mathbf{e} + \frac{\mathrm{m_e}}{\tau_\mathrm{e}}\mathbf{v_e} = -\mathrm{e}\mathbf{E} - \mathrm{e}(\mathbf{v_e}\times\mathbf{B}) \tag{B.4}$$

for electrons, and for holes

$$\mathrm{m_h}\dot{\mathbf{v}}_\mathbf{h} + \frac{\mathrm{m_h}}{\tau_\mathrm{h}}\mathbf{v_h} = \mathrm{e}\mathbf{E} + \mathrm{e}(\mathbf{v_h}\times\mathbf{B})\,. \tag{B.5}$$

For the stationary case ($\dot{\mathrm{v}} = 0$) and for electrons, Eq. B.4 simplifies by using the definition of the Drude mobility ($\mu = \mathrm{q}\tau/\mathrm{m}$) to

$$\begin{pmatrix} v_{e,x} \\ v_{e,y} \end{pmatrix} = -\mu_e \left[\begin{pmatrix} E_x \\ E_y \end{pmatrix} + \begin{pmatrix} v_{e,y}B \\ -v_{e,x}B \end{pmatrix} \right]. \tag{B.6}$$

The equation for $\mathrm{v_{e,y}}$ can be inserted in the equation for $\mathrm{v_{e,x}}$, and vice versa. For electrons this results in

$$\mathrm{v_{e,x}} = -\frac{\mu_\mathrm{e}(\mathrm{E_x} - \mu_\mathrm{e}\mathrm{BE_y})}{1+\mu_\mathrm{e}^2\mathrm{B}^2} \quad \text{and} \quad \mathrm{v_{e,y}} = -\frac{\mu_\mathrm{e}(\mathrm{E_y} + \mu_\mathrm{e}\mathrm{BE_x})}{1+\mu_\mathrm{e}^2\mathrm{B}^2}\,. \tag{B.7}$$

For holes we obtain

$$\mathrm{v_{h,x}} = \frac{\mu_\mathrm{h}(\mathrm{E_x} + \mu_\mathrm{h}\mathrm{BE_y})}{1+\mu_\mathrm{h}^2\mathrm{B}^2} \quad \text{and} \quad \mathrm{v_{h,y}} = \frac{\mu_\mathrm{h}(\mathrm{E_y} - \mu_\mathrm{h}\mathrm{BE_x})}{1+\mu_\mathrm{h}^2\mathrm{B}^2}\,. \tag{B.8}$$

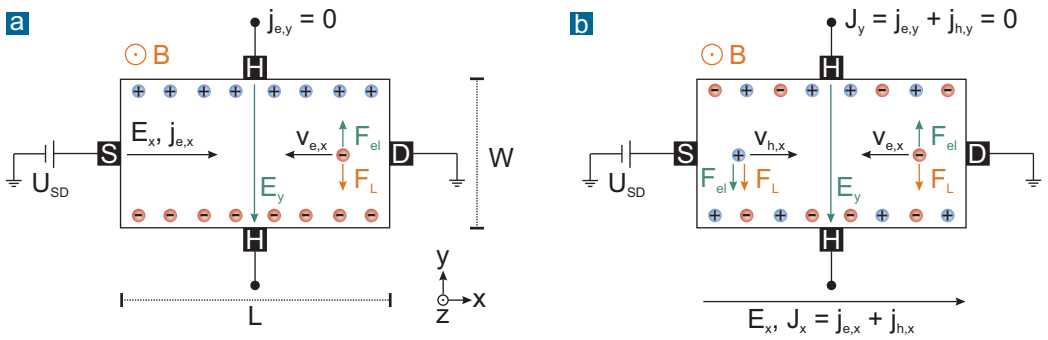

Figure B.2. Hall effect in a two-dimensional system. External fields and the corresponding forces on electrons, in a system with only electrons as charge carrier type, are shown in panel a. The case of coexisting electrons and holes is depicted in panel b. In both scenarios we consider a rectangular sample of length L and width W. The electrical current flows between the source (S) and drain (D) contacts while the Hall voltage is measured between the Hall contacts (H).

It is now possible to derive the expressions for the total current densities assuming an electron density n_e and a hole density n_h (see also Figure B.2b)

$$\begin{aligned} J_x &= j_{e,x} + j_{h,x} = -en_e v_{e,x} + en_h v_{h,x} \\ &= \frac{en_h\mu_h}{1+\mu_h^2 B^2}(E_x + \mu_h B E_y) + \frac{en_e\mu_e}{1+\mu_e^2 B^2}(E_x - \mu_e B E_y) \end{aligned} \tag{B.9}$$

and

$$\begin{aligned} J_y &= j_{e,y} + j_{h,y} = -en_e v_{e,y} + en_h v_{h,y} \\ &= \frac{en_e\mu_e}{1+\mu_e^2 B^2}(E_y + \mu_e B E_x) + \frac{en_h\mu_h}{1+\mu_h^2 B^2}(E_y - \mu_h B E_x) \end{aligned} \tag{B.10}$$

With the condition $J_y = 0$ these two equations can be rearranged and the Hall coefficient $R_H = E_y/J_x B$ can be calculated. Alternatively, the two equations can be merged to identify the elements of the conductivity tensor according to $\mathbf{J} = \hat{\sigma}\mathbf{E}$:

$$\mathbf{J} = \begin{pmatrix} \frac{en_h\mu_h}{1+\mu_h^2B^2} + \frac{en_e\mu_e}{1+\mu_e^2B^2} & \left(\frac{en_h\mu_h^2}{1+\mu_h^2B^2} - \frac{en_e\mu_e^2}{1+\mu_e^2B^2}\right) B \\ \left(\frac{en_e\mu_e^2}{1+\mu_e^2B^2} - \frac{en_h\mu_h^2}{1+\mu_h^2B^2}\right) B & \frac{en_h\mu_h}{1+\mu_h^2B^2} + \frac{en_e\mu_e}{1+\mu_e^2B^2} \end{pmatrix} \mathbf{E} \tag{B.11}$$

In the presence of a magnetic field the conductivity and resistivity acquire a tensorial form and can be translated into each other by $\hat{\rho} = \hat{\sigma}^{-1}$ to fulfill

$$\mathbf{J} = \hat{\sigma}\mathbf{E} \leftrightarrow \mathbf{E} = \hat{\rho}\mathbf{J}\,. \tag{B.12}$$

Due to the connection of the Lorentz force to the orientation of the magnetic field one can show that $\sigma_{xx} = \sigma_{yy}$ and $\sigma_{xy} = -\sigma_{yx}$, and so

$$\hat{\rho} = \begin{pmatrix} \rho_{xx} & \rho_{xy} \\ \rho_{yx} & \rho_{xx} \end{pmatrix} = \frac{1}{\sigma_{xx}^2 + \sigma_{xy}^2} \begin{pmatrix} \sigma_{xx} & -\sigma_{xy} \\ \sigma_{xy} & \sigma_{xx} \end{pmatrix}, \tag{B.13}$$

This yields the following expressions for the longitudinal and transverse resistivity

$$\rho_{xx} = \frac{\sigma_{xx}}{\sigma_{xx}^2 + \sigma_{xy}^2} \text{ and } \rho_{yx} = \frac{\sigma_{xy}}{\sigma_{xx}^2 + \sigma_{xy}^2}. \tag{B.14}$$

The transverse resistivity is then given by (using σ_{xx} and σ_{xy} from Eq. B.11)

$$\rho_{yx} = \frac{\frac{en_h\mu_h^2}{1+\mu_h^2B^2} - \frac{en_e\mu_e^2}{1+\mu_e^2B^2}}{\left(\frac{en_h\mu_h}{1+\mu_h^2B^2} + \frac{en_e\mu_e}{1+\mu_e^2B^2}\right)^2 + \left(\frac{en_h\mu_h^2}{1+\mu_h^2B^2} - \frac{en_e\mu_e^2}{1+\mu_e^2B^2}\right)^2 B^2} B \tag{B.15}$$

Since $E_y = \rho_{yx}J_x$ and $R_H \equiv E_y/J_xB$, the Hall coefficient for a two-dimensional system with coexisting electrons and holes is given by

$$R_H = \frac{\frac{en_h\mu_h^2}{1+\mu_h^2B^2} - \frac{en_e\mu_e^2}{1+\mu_e^2B^2}}{\left(\frac{en_h\mu_h}{1+\mu_h^2B^2} + \frac{en_e\mu_e}{1+\mu_e^2B^2}\right)^2 + \left(\frac{en_h\mu_h^2}{1+\mu_h^2B^2} - \frac{en_e\mu_e^2}{1+\mu_e^2B^2}\right)^2 B^2}. \tag{B.16}$$

As a consequence, the Hall coefficient R_H can possess a magnetic field dependence when $\mu_xB \geq 1$. Electron and hole mobilities in graphene are typically equal which reduces the above formula to

$$R_H = \frac{(n_h - n_e)(1 + \mu^2B^2)}{e[(n_h + n_e)^2 + (n_h - n_e)^2\mu^2B^2]}. \tag{B.17}$$

Accordingly, the longitudinal resistivity is given by

$$\rho_{xx} = \frac{\frac{en_h\mu_h}{1+\mu_h^2B^2} + \frac{en_e\mu_e}{1+\mu_e^2B^2}}{\left(\frac{en_h\mu_h}{1+\mu_h^2B^2} + \frac{en_e\mu_e}{1+\mu_e^2B^2}\right)^2 + \left(\frac{en_h\mu_h^2}{1+\mu_h^2B^2} - \frac{en_e\mu_e^2}{1+\mu_e^2B^2}\right)^2 B^2} \tag{B.18}$$

For equal electron and hole mobilities this reduces to

$$\rho_{xx} = \frac{(n_h + n_e)(1 + \mu^2B^2)}{e\mu[(n_h + n_e)^2 + (n_h - n_e)^2\mu^2B^2]}. \tag{B.19}$$

The presence of two types of charge carriers (or two electronic bands) gives rise to a non-zero classical magneto-resistance. This is also true for a spatially inhomogeneous

system like graphene in which electrons and holes break up into separated puddles with quasi-transparent boundaries at the charge neutrality point (see Section 2.3.3). If only one carrier type is present (single electronic band case), the resistance is field independent (see Section 2.3.1).

For weak magnetic fields ($\mu_x^2 B^2 \ll 1$), the longitudinal resistivity is given by

$$\rho_{xx}(B \approx 0) = \frac{1}{\sigma_{xx}} = \frac{1}{e(n_h \mu_h + n_e \mu_e)}, \tag{B.20}$$

and the Hall coefficient reduces to

$$R_H(B \approx 0) = \frac{n_h \mu_h^2 - n_e \mu_e^2}{e(n_h \mu_h + n_e \mu_e)^2} = \frac{n_h - b^2 n_e}{e(n_h + b n_e)^2}. \tag{B.21}$$

Here, $b = \mu_e/\mu_h$ refers to the ratio of the charge carrier mobilities.

So far we have assumed that all carriers have the same velocity (or energy). At finite temperature, the carrier energies follow the Fermi Dirac distribution f(E). The full calculation in the Boltzmann kinetic approach can be found in Ref. 150. The expression for the Hall coefficient then reads

$$R_H(B \approx 0) = \frac{n_h\, r_h\, \mu_h^2 - n_e\, r_e\, \mu_e^2}{e(n_h \mu_h + n_e \mu_e)^2}, \tag{B.22}$$

where $r_{e,h}$ is the Hall (scattering) factor for electrons and holes, respectively (see Section 2.3.1).

From the above formulas we can easily derive for a single charge carrier type by just setting one of the carrier densities to zero. For $n_h = 0$, $\rho_{xx}(B)$ follows from Eq. B.19:

$$\rho_{xx}(B) = \frac{1}{e\, n_e\, \mu_e} \quad (\text{indep. of } B), \tag{B.23}$$

and R_H from Eq. B.22:

$$R_H = -\frac{r_e}{e\, n_e}. \tag{B.24}$$

Bibliography

[1] K. S. Novoselov, A. K. Geim, S. V. Morozov, et al., "Electric field effect in atomically thin carbon films," *Science* **306**, 666–669 (2004).

[2] A. H. Castro Neto, F. Guinea, N. M. R. Peres, et al., "The electronic properties of graphene," *Reviews of Modern Physics* **81**, 109–162 (2009).

[3] S. Das Sarma, S. Adam, E. H. Hwang, et al., "Electronic transport in two-dimensional graphene," *Reviews of Modern Physics* **83**, 407–470 (2011).

[4] V. N. Kotov, B. Uchoa, V. M. Pereira, et al., "Electron-electron interactions in graphene: Current status and perspectives," *Reviews of Modern Physics* **84**, 1067–1125 (2012).

[5] D. R. Cooper, B. D'Anjou, N. Ghattamaneni, et al., "Experimental review of graphene," *ISRN Condensed Matter Physics* **2012**, 1–56 (2012).

[6] N. M. R. Peres, "Colloquium: The transport properties of graphene: An introduction," *Reviews of Modern Physics* **82**, 2673–2700 (2010).

[7] H. Raza, ed., *Graphene Nanoelectronics*. Springer, 1 ed. (2012).

[8] H. Aoki and M. Dresselhaus, eds., *Physics of Graphene*. Springer, 1 ed. (2014).

[9] V. Skakalova and A. Kaiser, eds., *Graphene*. Woodhead Publishing, 1 ed. (2014).

[10] L. Wang, I. Meric, P. Y. Huang, et al., "One-dimensional electrical contact to a two-dimensional material," *Science* **342**, 614–617 (2013).

[11] B. N. Narozhny, I. V. Gornyi, A. D. Mirlin, et al., "Hydrodynamic approach to electronic transport in graphene," *Annalen der Physik* **529**, 1700043 (2017).

[12] A. Lucas and K. Chung Fong, "Hydrodynamics of electrons in graphene," *Journal of Physics: Condensed Matter* **30**, 053001 (2018).

[13] C. Lee, X. Wei, J. W. Kysar, et al., "Measurement of the elastic properties and intrinsic strength of monolayer graphene," *Science* **321**, 385–388 (2008).

[14] K. S. Novoselov, D. Jiang, F. Schedin, et al., "Two-dimensional atomic crystals," *Proceedings of the National Academy of Sciences of the United States of America* **102**, 10451–10453 (2005).

[15] A. K. Geim and I. V. Grigorieva, "Van der waals heterostructures," *Nature* **499**, 419–425 (2013).

[16] J. C. Meyer, A. K. Geim, M. I. Katsnelson, et al., "The structure of suspended graphene sheets," *Nature* **446**, 60–63 (2007).

[17] A. Fasolino, J. H. Los, and M. I. Katsnelson, "Intrinsic ripples in graphene," *Nature Materials* **6**, 858–861 (2007).

[18] N. Couto, D. Costanzo, S. Engels, et al., "Random strain fluctuations as dominant disorder source for high-quality on-substrate graphene devices," *Physical Review X* **4**, 041019 (2014).

[19] B. Krauß, *Scanning probe spectroscopy of graphene nanostructures.* PhD thesis, University of Stuttgart (2012).

[20] P. R. Wallace, "The band theory of graphite," *Physical Review* **71**, 622–634 (1947).

[21] R. S. Deacon, K. C. Chuang, R. J. Nicholas, et al., "Cyclotron resonance study of the electron and hole velocity in graphene monolayers," *Physical Review B* **76**, 081406 (2007).

[22] D. H. Chae, T. Utikal, S. Weisenburger, et al., "Excitonic fano resonance in free-standing graphene," *Nano Letters* **11**, 1379–1382 (2011).

[23] A. Bostwick, T. Ohta, T. Seyller, et al., "Quasiparticle dynamics in graphene," *Nature Physics* **3**, 36–40 (2006).

[24] S. Y. Zhou, G. H. Gweon, J. Graf, et al., "First direct observation of dirac fermions in graphite," *Nature Physics* **2**, 595–599 (2006).

[25] T. Ando, T. Nakanishi, and R. Saito, "Berry's phase and absence of back scattering in carbon nanotubes," *Journal of the Physical Society of Japan* **67**, 2857–2862 (1998).

[26] P. L. McEuen, M. Bockrath, D. H. Cobden, et al., "Disorder, pseudospins, and backscattering in carbon nanotubes," *Physical Review Letters* **83**, 5098–5101 (1999).

[27] Y. Zhang, Y. W. Tan, H. L. Stormer, et al., "Experimental observation of the quantum hall effect and berry's phase in graphene," *Nature* **438**, 201–204 (2005).

[28] T. Ando, "Screening effect and impurity scattering in monolayer graphene," *Journal of the Physical Society of Japan* **75**, 074716 (2006).

[29] S. Adam, E. H. Hwang, E. Rossi, et al., "Theory of charged impurity scattering in two-dimensional graphene," *Solid State Communications* **149**, 1072–1079 (2009).

[30] C. Jang, S. Adam, J. H. Chen, et al., "Tuning the effective fine structure constant in graphene: opposing effects of dielectric screening on short- and long-range potential scattering," *Physical Review Letters* **101**, 146805 (2008).

[31] S. Adam, S. Jung, N. N. Klimov, et al., "Mechanism for puddle formation in graphene," *Physical Review B* **84**, 235421 (2011).

[32] E. H. Hwang, S. Adam, and S. D. Sarma, "Carrier transport in two-dimensional graphene layers," *Physical Review Letters* **98**, 186806 (2007).

[33] K. S. Novoselov, A. K. Geim, S. V. Morozov, et al., "Two-dimensional gas of massless dirac fermions in graphene," *Nature* **438**, 197–200 (2005).

[34] E. McCann and M. Koshino, "The electronic properties of bilayer graphene," *Reports on Progress in Physics* **76**, 056503 (2013).

[35] A. V. Rozhkov, A. O. Sboychakov, A. L. Rakhmanov, et al., "Electronic properties of graphene-based bilayer systems," *Physics Reports* **648**, 1–104 (2016).

[36] C. Ortix, L. Yang, and J. van den Brink, "Graphene on incommensurate substrates: Trigonal warping and emerging dirac cone replicas with halved group velocity," *Physical Review B* **86**, 081405 (2012).

[37] M. I. Katsnelson, K. S. Novoselov, and A. K. Geim, "Chiral tunnelling and the klein paradox in graphene," *Nature Physics* **2**, 620–625 (2006).

[38] S. Adam, E. H. Hwang, and S. Das Sarma, "Scattering mechanisms and boltzmann transport in graphene," *Physica E: Low-dimensional Systems and Nanostructures* **40**, 1022–1025 (2008).

[39] S. V. Morozov, K. S. Novoselov, M. I. Katsnelson, et al., "Giant intrinsic carrier mobilities in graphene and its bilayer," *Physical Review Letters* **100**, 016602 (2008).

[40] S. Engels, *Limitations to charge carrier transport in high quality single-and bi-layer graphene.* PhD thesis, RWTH Aachen University (2015).

[41] E. McCann, "Staying or going? chirality decides!," *Physics* **2**, 98 (2009).

[42] E. McCann, K. Kechedzhi, V. I. Fal'ko, et al., "Weak-localization magnetoresistance and valley symmetry in graphene," *Physical Review Letters* **97**, 146805 (2006).

[43] Y. Kim, P. Herlinger, P. Moon, et al., "Charge inversion and topological phase transition at a twist angle induced van hove singularity of bilayer graphene," *Nano Letters* **16**, 5053–5059 (2016).

[44] M. Yankowitz, J. Xue, D. Cormode, et al., "Emergence of superlattice dirac points in graphene on hexagonal boron nitride," *Nature Physics* **8**, 382–386 (2012).

[45] L. A. Ponomarenko, R. V. Gorbachev, G. L. Yu, et al., "Cloning of dirac fermions in graphene superlattices," *Nature* **497**, 594–597 (2013).

[46] A. M. DaSilva, J. Jung, S. Adam, et al., "Transport and particle-hole asymmetry in graphene on boron nitride," *Physical Review B* **91**, 245422 (2015).

[47] L. Wang, Y. Gao, B. Wen, et al., "Evidence for a fractional fractal quantum hall effect in graphene superlattices," *Science* **350**, 1231–1234 (2015).

[48] R. Ribeiro-Palau, C. Zhang, K. Watanabe, et al., "Twistable electronics with dynamically rotatable heterostructures." arXiv:1804.02038 [cond-mat.mes-hall].

[49] C. R. Woods, L. Britnell, A. Eckmann, et al., "Commensurate–incommensurate transition in graphene on hexagonal boron nitride," *Nature Physics* **10**, 451–456 (2014).

[50] C. R. Dean, L. Wang, P. Maher, et al., "Hofstadter's butterfly and the fractal quantum hall effect in moire superlattices," *Nature* **497**, 598–602 (2013).

[51] J. Jung, A. M. DaSilva, A. H. MacDonald, et al., "Origin of band gaps in graphene on hexagonal boron nitride," *Nature Communications* **6**, 6308 (2015).

[52] B. Hunt, J. D. Sanchez-Yamagishi, A. F. Young, et al., "Massive dirac fermions and hofstadter butterfly in a van der waals heterostructure," *Science* **340**, 1427–1430 (2013).

[53] M. Yankowitz, J. Jung, E. Laksono, et al., "Dynamic band structure tuning of graphene moir\'e superlattices with pressure," *Nature* **557**, 404–408 (2018).

[54] M. A. H. Vozmediano, "Graphene: The running of the constants," *Nature Physics* **7**, 671–672 (2011).

[55] J. Chae, S. Jung, A. F. Young, et al., "Renormalization of the graphene dispersion velocity determined from scanning tunneling spectroscopy," *Physical Review Letters* **109**, 116802 (2012).

[56] D. C. Elias, R. V. Gorbachev, A. S. Mayorov, et al., "Dirac cones reshaped by interaction effects in suspended graphene," *Nature Physics* **7**, 701–704 (2011).

[57] S. Das Sarma, E. H. Hwang, and W.-K. Tse, "Many-body interaction effects in doped and undoped graphene: Fermi liquid versus non-fermi liquid," *Physical Review B* **75**, 121406 (2007).

[58] S. Das Sarma and E. H. Hwang, "Velocity renormalization and anomalous quasiparticle dispersion in extrinsic graphene," *Physical Review B* **87**, 045425 (2013).

[59] J. Hofmann, E. Barnes, and S. Das Sarma, "Why does graphene behave as a weakly interacting system?," *Physical Review Letters* **113**, 105502 (2014).

[60] C. Bauer, A. Rückriegel, A. Sharma, et al., "Nonperturbative renormalization group calculation of quasiparticle velocity and dielectric function of graphene," *Physical Review B* **92**, 121409 (2015).

[61] A. Sharma and P. Kopietz, "Multilogarithmic velocity renormalization in graphene," *Physical Review B* **93**, 235425 (2016).

[62] T. Stauber, P. Parida, M. Trushin, et al., "Interacting electrons in graphene: Fermi velocity renormalization and optical response," *Physical Review Letters* **118**, 266801 (2017).

[63] D. A. Siegel, W. Regan, A. V. Fedorov, et al., "Charge-carrier screening in single-layer graphene," *Physical Review Letters* **110**, 146802 (2013).

[64] C. Kumar, S. K. Srivastav, P. Adhikary, et al., "Localization physics in graphene Moire superlattices." arXiv:1711.03275 [cond-mat.mes-hall].

[65] Z. Q. Li, E. A. Henriksen, Z. Jiang, et al., "Dirac charge dynamics in graphene by infrared spectroscopy," *Nature Physics* **4**, 532–535 (2008).

[66] A. Luican, G. Li, and E. Y. Andrei, "Quantized landau level spectrum and its density dependence in graphene," *Physical Review B* **83**, 041405 (2011).

[67] C. Faugeras, S. Berciaud, P. Leszczynski, et al., "Landau level spectroscopy of electron-electron interactions in graphene," *Physical Review Letters* **114**, 126804 (2015).

[68] B. Terres, L. A. Chizhova, F. Libisch, et al., "Size quantization of dirac fermions in graphene constrictions," *Nature Communications* **7**, 11528 (2016).

[69] D. A. Siegel, C. H. Park, C. Hwang, et al., "Many-body interactions in quasi-freestanding graphene," *Proceedings of the National Academy of Sciences of the United States of America* **108**, 11365–11369 (2011).

[70] C. Hwang, D. A. Siegel, S.-K. Mo, et al., "Fermi velocity engineering in graphene by substrate modification," *Scientific Reports* **2**, 590 (2012).

[71] G. L. Yu, R. Jalil, B. Belle, et al., "Interaction phenomena in graphene seen through quantum capacitance," *Proceedings of the National Academy of Sciences of the United States of America* **110**, 3282–3286 (2013).

[72] K. I. Bolotin, F. Ghahari, M. D. Shulman, et al., "Observation of the fractional quantum hall effect in graphene," *Nature* **462**, 196–199 (2009).

[73] X. Du, I. Skachko, F. Duerr, et al., "Fractional quantum hall effect and insulating phase of dirac electrons in graphene," *Nature* **462**, 192–195 (2009).

[74] C. R. Dean, A. F. Young, P. Cadden-Zimansky, et al., "Multicomponent fractional quantum hall effect in graphene," *Nature Physics* **7**, 693–696 (2011).

[75] D. N. Basov and M. M. Fogler, "Plasmonic imaging is gaining momentum," *Science* **357**, 132–133 (2017).

[76] J. Jobst, D. Waldmann, I. V. Gornyi, et al., "Electron-electron interaction in the magnetoresistance of graphene," *Physical Review Letters* **108**, 106601 (2012).

[77] F. V. Tikhonenko, A. A. Kozikov, A. K. Savchenko, et al., "Transition between electron localization and antilocalization in graphene," *Physical Review Letters* **103**, 226801 (2009).

[78] T. Sohier, M. Calandra, C.-H. Park, et al., "Phonon-limited resistivity of graphene by first-principles calculations: Electron-phonon interactions, strain-induced gauge field, and boltzmann equation," *Physical Review B* **90**, 125414 (2014).

[79] J. Crossno, J. K. Shi, K. Wang, et al., "Observation of the dirac fluid and the breakdown of the wiedemann-franz law in graphene," *Science* **351**, 1058–1061 (2016).

[80] D. Y. H. Ho, I. Yudhistira, N. Chakraborty, et al., "Theoretical determination of hydrodynamic window in monolayer and bilayer graphene from scattering rates," *Physical Review B* **97**, 121404 (2018).

[81] Y. Nam, D.-K. Ki, D. Soler-Delgado, et al., "Electron–hole collision limited transport in charge-neutral bilayer graphene," *Nature Physics* **13**, 1207–1214 (2017).

[82] W. Zhu, V. Perebeinos, M. Freitag, et al., "Carrier scattering, mobilities, and electrostatic potential in monolayer, bilayer, and trilayer graphene," *Physical Review B* **80**, 235402 (2009).

[83] P. Blake, E. W. Hill, A. H. Castro Neto, et al., "Making graphene visible," *Applied Physics Letters* **91**, 063124 (2007).

[84] F. Paolucci, *Transport properties of electrolyte gated graphene devices*. PhD thesis, University of Stuttgart (2015).

[85] J. Ye, M. F. Craciun, M. Koshino, et al., "Accessing the transport properties of graphene and its multilayers at high carrier density," *Proceedings of the National Academy of Sciences of the United States of America* **108**, 13002–13006 (2011).

[86] A. Das, S. Pisana, B. Chakraborty, et al., "Monitoring dopants by raman scattering in an electrochemically top-gated graphene transistor," *Nature Nanotechnology* **3**, 210–215 (2008).

[87] D. K. Efetov and P. Kim, "Controlling electron-phonon interactions in graphene at ultrahigh carrier densities," *Physical Review Letters* **105**, 256805 (2010).

[88] M. Kühne, F. Paolucci, J. Popovic, et al., "Ultrafast lithium diffusion in bilayer graphene," *Nature Nanotechnology* **12**, 895–900 (2017).

[89] S. Y. F. Zhao, G. A. Elbaz, D. Kwabena Bediako, et al., "Controlled Electrochemical Intercalation of Graphene/h-BN van der Waals Heterostructures," *Nano Letters* **18**, 460–466 (2018).

[90] J. H. Chen, C. Jang, S. Xiao, et al., "Intrinsic and extrinsic performance limits of graphene devices on sio2," *Nature Nanotechnology* **3**, 206–209 (2008).

[91] Y. W. Tan, Y. Zhang, K. Bolotin, et al., "Measurement of scattering rate and minimum conductivity in graphene," *Physical Review Letters* **99**, 246803 (2007).

[92] C. Caillier, D. K. Ki, Y. Lisunova, et al., "Identification of a strong contamination source for graphene in vacuum systems," *Nanotechnology* **24**, 405201 (2013).

[93] J. A. Alexander-Webber, A. A. Sagade, A. I. Aria, et al., "Encapsulation of graphene transistors and vertical device integration by interface engineering with atomic layer deposited oxide," *2D Materials* **4**, 011008 (2016).

[94] K. Nagashio, T. Yamashita, T. Nishimura, et al., "Electrical transport properties of graphene on sio2 with specific surface structures," *Journal of Applied Physics* **110**, 024513 (2011).

[95] S. Ryu, L. Liu, S. Berciaud, et al., "Atmospheric oxygen binding and hole doping in deformed graphene on a sio(2) substrate," *Nano Letters* **10**, 4944–4951 (2010).

[96] J. Martin, N. Akerman, G. Ulbricht, et al., "Observation of electron–hole puddles in graphene using a scanning single-electron transistor," *Nature Physics* **4**, 144–148 (2007).

[97] Q. Li, E. H. Hwang, and S. Das Sarma, "Disorder-induced temperature-dependent transport in graphene: Puddles, impurities, activation, and diffusion," *Physical Review B* **84**, 115442 (2011).

[98] S. Samaddar, I. Yudhistira, S. Adam, et al., "Charge puddles in graphene near the dirac point," *Physical Review Letters* **116**, 126804 (2016).

[99] S. Adam, E. H. Hwang, V. M. Galitski, et al., "A self-consistent theory for graphene transport," *Proceedings of the National Academy of Sciences of the United States of America* **104**, 18392–18397 (2007).

[100] E. H. Hwang and S. Das Sarma, "Dielectric function, screening, and plasmons in two-dimensional graphene," *Physical Review B* **75**, 205418 (2007).

[101] B. Wunsch, T. Stauber, F. Sols, et al., "Dynamical polarization of graphene at finite doping," *New Journal of Physics* **8**, 318–318 (2006).

[102] E. H. Hwang and S. Das Sarma, "Screening-induced temperature-dependent transport in two-dimensional graphene," *Physical Review B* **79**, 165404 (2009).

[103] D. Wong, F. Corsetti, Y. Wang, et al., "Spatially resolving density-dependent screening around a single charged atom in graphene," *Physical Review B* **95**, 205419 (2017).

[104] J. Xue, J. Sanchez-Yamagishi, D. Bulmash, et al., "Scanning tunnelling microscopy and spectroscopy of ultra-flat graphene on hexagonal boron nitride," *Nature Materials* **10**, 282–285 (2011).

[105] E. Rossi and S. Das Sarma, "Ground state of graphene in the presence of random charged impurities," *Physical Review Letters* **101**, 166803 (2008).

[106] E. Rossi, S. Adam, and S. Das Sarma, "Effective medium theory for disordered two-dimensional graphene," *Physical Review B* **79**, 245423 (2009).

[107] M. Polini, A. Tomadin, R. Asgari, et al., "Density functional theory of graphene sheets," *Physical Review B* **78**, 115426 (2008).

[108] M. Gibertini, A. Tomadin, F. Guinea, et al., "Electron-hole puddles in the absence of charged impurities," *Physical Review B* **85**, 201405 (2012).

[109] S. Adam and M. D. Stiles, "Temperature dependence of the diffusive conductivity of bilayer graphene," *Physical Review B* **82**, 075423 (2010).

[110] J. Ping, I. Yudhistira, N. Ramakrishnan, et al., "Disorder-induced magnetoresistance in a two-dimensional electron system," *Physical Review Letters* **113**, 047206 (2014).

[111] E. H. Hwang and S. Das Sarma, "Insulating behavior in metallic bilayer graphene: Interplay between density inhomogeneity and temperature," *Physical Review B* **82**, 081409 (2010).

[112] I. Meric, M. Y. Han, A. F. Young, et al., "Current saturation in zero-bandgap, top-gated graphene field-effect transistors," *Nature Nanotechnology* **3**, 654–659 (2008).

[113] S. Kim, J. Nah, I. Jo, et al., "Realization of a high mobility dual-gated graphene field-effect transistor with al2o3 dielectric," *Applied Physics Letters* **94**, 062107 (2009).

[114] B. Chen, L. Huang, X. Ma, et al., "Exploration of sensitivity limit for graphene magnetic sensors," *Carbon* **94**, 585–589 (2015).

[115] Q. Li, *Theory of graphene transport properties*. PhD thesis, University of Maryland (2013).

[116] A. Deshpande, W. Bao, F. Miao, et al., "Spatially resolved spectroscopy of monolayer graphene on sio2," *Physical Review B* **79**, 205411 (2009).

[117] M. L. Teague, A. P. Lai, J. Velasco, et al., "Evidence for strain-induced local conductance modulations in single-layer graphene on sio2," *Nano Letters* **9**, 2542–2546 (2009).

[118] Y. Zhang, V. W. Brar, C. Girit, et al., "Origin of spatial charge inhomogeneity in graphene," *Nature Physics* **5**, 722–726 (2009).

[119] A. Deshpande, W. Bao, Z. Zhao, et al., "Imaging charge density fluctuations in graphene using coulomb blockade spectroscopy," *Physical Review B* **83**, 155409 (2011).

[120] R. Decker, Y. Wang, V. W. Brar, et al., "Local electronic properties of graphene on a bn substrate via scanning tunneling microscopy," *Nano Letters* **11**, 2291–2295 (2011).

[121] K. M. Burson, W. G. Cullen, S. Adam, et al., "Direct imaging of charged impurity density in common graphene substrates," *Nano Letters* **13**, 3576–3580 (2013).

[122] A. Woessner, P. Alonso-Gonzalez, M. B. Lundeberg, et al., "Near-field photocurrent nanoscopy on bare and encapsulated graphene," *Nature Communications* **7**, 10783 (2016).

[123] C. Neumann, S. Reichardt, P. Venezuela, et al., "Raman spectroscopy as probe of nanometre-scale strain variations in graphene," *Nature Communications* **6**, 8429 (2015).

[124] L. Banszerus, H. Janssen, M. Otto, et al., "Identifying suitable substrates for high-quality graphene-based heterostructures," *2D Materials* **4**, 025030 (2017).

[125] J. D. Thomsen, T. Gunst, S. S. Gregersen, et al., "Suppression of intrinsic roughness in encapsulated graphene," *Physical Review B* **96**, 014101 (2017).

[126] A. S. Mayorov, D. C. Elias, I. S. Mukhin, et al., "How close can one approach the dirac point in graphene experimentally?," *Nano Letters* **12**, 4629–4634 (2012).

[127] V. Geringer, M. Liebmann, T. Echtermeyer, et al., "Intrinsic and extrinsic corrugation of monolayer graphene deposited on sio2," *Physical Review Letters* **102**, 076102 (2009).

[128] W. G. Cullen, M. Yamamoto, K. M. Burson, et al., "High-fidelity conformation of graphene to sio2 topographic features," *Physical Review Letters* **105**, 215504 (2010).

[129] S. J. Haigh, A. Gholinia, R. Jalil, et al., "Cross-sectional imaging of individual layers and buried interfaces of graphene-based heterostructures and superlattices," *Nature Materials* **11**, 764–767 (2012).

[130] E. V. Castro, H. Ochoa, M. I. Katsnelson, et al., "Limits on charge carrier mobility in suspended graphene due to flexural phonons," *Physical Review Letters* **105**, 266601 (2010).

[131] B. Amorim, A. Cortijo, F. de Juan, et al., "Novel effects of strains in graphene and other two dimensional materials," *Physics Reports* **617**, 1–54 (2016).

[132] P. Partovi-Azar, N. Nafari, and M. R. R. Tabar, "Interplay between geometrical structure and electronic properties in rippled free-standing graphene," *Physical Review B* **83**, 165434 (2011).

[133] M. I. Katsnelson and A. K. Geim, "Electron scattering on microscopic corrugations in graphene," *Philosophical transactions Series A, Mathematical, physical, and engineering sciences* **366**, 195–204 (2008).

[134] C. H. Lui, L. Liu, K. F. Mak, et al., "Ultraflat graphene," *Nature* **462**, 339–341 (2009).

[135] L. A. Ponomarenko, R. Yang, T. M. Mohiuddin, et al., "Effect of a high-kappa environment on charge carrier mobility in graphene," *Physical Review Letters* **102**, 206603 (2009).

[136] M. Yamamoto, T. L. Einstein, M. S. Fuhrer, et al., "Charge inhomogeneity determines oxidative reactivity of graphene on substrates," *ACS Nano* **6**, 8335–8341 (2012).

[137] M. Yankowitz, S. Larentis, K. Kim, et al., "Intrinsic disorder in graphene on transition metal dichalcogenide heterostructures," *Nano Letters* **15**, 1925–1929 (2015).

[138] A. V. Kretinin, Y. Cao, J. S. Tu, et al., "Electronic properties of graphene encapsulated with different two-dimensional atomic crystals," *Nano Letters* **14**, 3270–3276 (2014).

[139] Q. H. Wang, Z. Jin, K. K. Kim, et al., "Understanding and controlling the substrate effect on graphene electron-transfer chemistry via reactivity imprint lithography," *Nature Chemistry* **4**, 724–732 (2012).

[140] T. Fang, A. Konar, H. Xing, et al., "Carrier statistics and quantum capacitance of graphene sheets and ribbons," *Applied Physics Letters* **91**, 092109 (2007).

[141] J. Fernández-Rossier, J. J. Palacios, and L. Brey, "Electronic structure of gated graphene and graphene ribbons," *Physical Review B* **75**, 205441 (2007).

[142] P. Wehrfritz and T. Seyller, "The hall coefficient: a tool for characterizing graphene field effect transistors," *2D Materials* **1**, 035004 (2014).

[143] L. Wang, W. Wang, G. Xu, et al., "Analytical carrier density and quantum capacitance for graphene," *Applied Physics Letters* **108**, 013503 (2016).

[144] H. Xu, Z. Zhang, Z. Wang, et al., "Quantum capacitance limited vertical scaling of graphene field-effect transistor," *ACS Nano* **5**, 2340–2347 (2011).

[145] S. Dröscher, P. Roulleau, F. Molitor, et al., "Quantum capacitance and density of states of graphene," *Applied Physics Letters* **96**, 152104 (2010).

[146] K. Nagashio, T. Nishimura, and A. Toriumi, "Estimation of residual carrier density near the dirac point in graphene through quantum capacitance measurement," *Applied Physics Letters* **102**, 173507 (2013).

[147] J. Xia, F. Chen, J. Li, et al., "Measurement of the quantum capacitance of graphene," *Nature Nanotechnology* **4**, 505–509 (2009).

[148] G. L. Yu, R. V. Gorbachev, J. S. Tu, et al., "Hierarchy of hofstadter states and replica quantum hall ferromagnetism in graphene superlattices," *Nature Physics* **10**, 525–529 (2014).

[149] S. Das Sarma, E. H. Hwang, and Q. Li, "Disorder by order in graphene," *Physical Review B* **85**, 195451 (2012).

[150] R. S. Popovic, *Hall Effect Devices*. IoP Publishing, 2 ed. (2003).

[151] T. Ihn, *Semiconductor Nanostructures*. Oxford University Press, 1 ed. (2010).

[152] P. Yu and M. Cardona, *Fundamentals of Semiconductors*. Springer, 4 ed. (2010).

[153] C. Herring and E. Vogt, "Transport and deformation-potential theory for many-valley semiconductors with anisotropic scattering," *Physical Review* **101**, 944–961 (1956).

[154] A. Asgari, S. Babanejad, and L. Faraone, "Electron mobility, hall scattering factor, and sheet conductivity in algan/aln/gan heterostructures," *Journal of Applied Physics* **110**, 113713 (2011).

[155] D. Rode, "Theory of electron galvanomagnetics in crystals: Hall effect in semiconductors and semimetals," *physica status solidi b* **55**, 687–696 (1973).

[156] G. Ng, D. Vasileska, and D. K. Schroder, "Calculation of the electron hall mobility and hall scattering factor in 6h-sic," *Journal of Applied Physics* **106**, 053719 (2009).

[157] H. Iwata and K. M. Itoh, "Donor and acceptor concentration dependence of the electron hall mobility and the hall scattering factor in n-type 4h– and 6h–sic," *Journal of Applied Physics* **89**, 6228–6234 (2001).

[158] S. Cho and M. S. Fuhrer, "Charge transport and inhomogeneity near the minimum conductivity point in graphene," *Physical Review B* **77**, 081402 (2008).

[159] H. Ibach and H. Lüth, *Festkörperphysik*. Springer, 7 ed. (2009).

[160] S. Das Sarma and E. H. Hwang, "Density-dependent electrical conductivity in suspended graphene: Approaching the dirac point in transport," *Physical Review B* **87**, 035415 (2013).

[161] C. R. Dean, A. F. Young, I. Meric, et al., "Boron nitride substrates for high-quality graphene electronics," *Nature Nanotechnology* **5**, 722–726 (2010).

[162] S. Masubuchi, K. Iguchi, T. Yamaguchi, et al., "Boundary scattering in ballistic graphene," *Physical Review Letters* **109**, 036601 (2012).

[163] J. H. Chen, C. Jang, M. Ishigami, et al., "Diffusive charge transport in graphene on sio2," *Solid State Communications* **149**, 1080–1086 (2009).

[164] A. W. Tsen, L. Brown, M. P. Levendorf, et al., "Tailoring electrical transport across grain boundaries in polycrystalline graphene," *Science* **336**, 1143–1146 (2012).

[165] T. Stauber, N. M. R. Peres, and F. Guinea, "Electronic transport in graphene: A semiclassical approach including midgap states," *Physical Review B* **76**, 205423 (2007).

[166] E. R. Mucciolo and C. H. Lewenkopf, "Disorder and electronic transport in graphene," *Journal of Physics Condensed Matter* **22**, 273201 (2010).

[167] A. Ferreira, J. Viana-Gomes, J. Nilsson, et al., "Unified description of the dc conductivity of monolayer and bilayer graphene at finite densities based on resonant scatterers," *Physical Review B* **83**, 165402 (2011).

[168] Z. H. Ni, L. A. Ponomarenko, R. R. Nair, et al., "On resonant scatterers as a factor limiting carrier mobility in graphene," *Nano Letters* **10**, 3868–3872 (2010).

[169] S. Das Sarma and Q. Li, "Graphene on srtio3," *Solid State Communications* **152**, 1795–1799 (2012).

[170] D. Halbertal, M. Ben Shalom, A. Uri, et al., "Imaging resonant dissipation from individual atomic defects in graphene," *Science* **358**, 1303–1306 (2017).

[171] J. H. Chen, W. G. Cullen, C. Jang, et al., "Defect scattering in graphene," *Physical Review Letters* **102**, 236805 (2009).

[172] J. Katoch, J. H. Chen, R. Tsuchikawa, et al., "Uncovering the dominant scatterer in graphene sheets on sio2," *Physical Review B* **82**, 081417 (2010).

[173] N. J. Couto, B. Sacepe, and A. F. Morpurgo, "Transport through graphene on srtio3," *Physical Review Letters* **107**, 225501 (2011).

[174] Q. Li, E. H. Hwang, E. Rossi, et al., "Theory of 2d transport in graphene for correlated disorder," *Physical Review Letters* **107**, 156601 (2011).

[175] E. H. Hwang and S. Das Sarma, "Acoustic phonon scattering limited carrier mobility in two-dimensional extrinsic graphene," *Physical Review B* **77**, 115449 (2008).

[176] K. Zou, X. Hong, D. Keefer, et al., "Deposition of high-quality hfo2 on graphene and the effect of remote oxide phonon scattering," *Physical Review Letters* **105**, 126601 (2010).

[177] A. K. Newaz, Y. S. Puzyrev, B. Wang, et al., "Probing charge scattering mechanisms in suspended graphene by varying its dielectric environment," *Nature Communications* **3**, 734 (2012).

[178] J. H. Chen, C. Jang, S. Adam, et al., "Charged-impurity scattering in graphene," *Nature Physics* **4**, 377–381 (2008).

[179] S. Xiao, J.-H. Chen, S. Adam, et al., "Charged impurity scattering in bilayer graphene," *Physical Review B* **82**, 041406 (2010).

[180] S. Das Sarma and E. H. Hwang, "Conductivity of graphene on boron nitride substrates," *Physical Review B* **83**, 121405 (2011).

[181] M. A. H. Vozmediano, M. I. Katsnelson, and F. Guinea, "Gauge fields in graphene," *Physics Reports* **496**, 109–148 (2010).

[182] M. Ishigami, J. H. Chen, W. G. Cullen, et al., "Atomic structure of graphene on sio2," *Nano Letters* **7**, 1643–1648 (2007).

[183] Y. Kim, P. Herlinger, T. Taniguchi, et al., "Reliable postprocessing improvement of van der waals heterostructures," *ACS Nano* **13**, 14182–14190 (2019).

[184] L. Wang, P. Makk, S. Zihlmann, et al., "Mobility enhancement in graphene by in situ reduction of random strain fluctuations," *Physical Review Letters* **124**, 157701 (2020).

[185] S. Zihlmann, P. Makk, M. K. Rehmann, et al., "Out-of-plane corrugations in graphene based van der waals heterostructures," *Physical Review B* **102**, 195404 (2020).

[186] J. Yan and M. S. Fuhrer, "Correlated charged impurity scattering in graphene," *Physical Review Letters* **107**, 206601 (2011).

[187] M. Fuhrer, "Textbook physics from a cutting-edge material," *Physics* **3**, 106 (2010).

[188] F. Ghahari, H. Y. Xie, T. Taniguchi, et al., "Enhanced thermoelectric power in graphene: Violation of the mott relation by inelastic scattering," *Physical Review Letters* **116**, 136802 (2016).

[189] S. Fratini and F. Guinea, "Substrate-limited electron dynamics in graphene," *Physical Review B* **77**, 195415 (2008).

[190] J. Schiefele, F. Sols, and F. Guinea, "Temperature dependence of the conductivity of graphene on boron nitride," *Physical Review B* **85**, 195420 (2012).

[191] A. Konar, T. Fang, and D. Jena, "Effect of high-k gate dielectrics on charge transport in graphene-based field effect transistors," *Physical Review B* **82**, 115452 (2010).

[192] H.-Y. Xie and M. S. Foster, "Transport coefficients of graphene: Interplay of impurity scattering, coulomb interaction, and optical phonons," *Physical Review B* **93**, 195103 (2016).

[193] S. Wiedmann, H. J. van Elferen, E. V. Kurganova, et al., "Coexistence of electron and hole transport in graphene," *Physical Review B* **84**, 115314 (2011).

[194] M. M. Fogler, "Neutrality point of graphene with coplanar charged impurities," *Physical Review Letters* **103**, 236801 (2009).

[195] A. S. Mayorov, R. V. Gorbachev, S. V. Morozov, et al., "Micrometer-scale ballistic transport in encapsulated graphene at room temperature," *Nano Letters* **11**, 2396–2399 (2011).

[196] X. Du, I. Skachko, A. Barker, et al., "Approaching ballistic transport in suspended graphene," *Nature Nanotechnology* **3**, 491–495 (2008).

[197] Z. Hu, D. Prasad Sinha, J. Ung Lee, et al., "Substrate dielectric effects on graphene field effect transistors," *Journal of Applied Physics* **115**, 194507 (2014).

[198] M. K. Joo, J. Kim, J. H. Park, et al., "Large-scale graphene on hexagonal-bn hall elements: Prediction of sensor performance without magnetic field," *ACS Nano* **10**, 8803–11 (2016).

[199] V. Guttal and D. Stroud, "Model for a macroscopically disordered conductor with an exactly linear high-field magnetoresistance," *Physical Review B* **71**, 201304 (2005).

[200] L. A. Ponomarenko, A. K. Geim, A. A. Zhukov, et al., "Tunable metal–insulator transition in double-layer graphene heterostructures," *Nature Physics* **7**, 958–961 (2011).

[201] F. Amet, J. R. Williams, K. Watanabe, et al., "Insulating behavior at the neutrality point in single-layer graphene," *Physical Review Letters* **110**, 216601 (2013).

[202] M. Rodriguez-Vega, J. Fischer, S. Das Sarma, et al., "Ground state of graphene heterostructures in the presence of random charged impurities," *Physical Review B* **90**, 035406 (2014).

[203] L. Banszerus, M. Schmitz, S. Engels, et al., "Ballistic transport exceeding 28 mum in cvd grown graphene," *Nano Letters* **16**, 1387–1391 (2016).

[204] K. I. Bolotin, K. J. Sikes, J. Hone, et al., "Temperature-dependent transport in suspended graphene," *Physical Review Letters* **101**, 096802 (2008).

[205] R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, et al., "Superballistic flow of viscous electron fluid through graphene constrictions," *Nature Physics* **13**, 1182–1185 (2017).

[206] S. Chen, Z. Han, M. M. Elahi, et al., "Electron optics with p-n junctions in ballistic graphene," *Science* **353**, 1522–1525 (2016).

[207] A. W. Barnard, A. Hughes, A. L. Sharpe, et al., "Absorptive pinhole collimators for ballistic dirac fermions in graphene," *Nature Communications* **8**, 15418 (2017).

[208] G.-H. Lee, G.-H. Park, and H.-J. Lee, "Observation of negative refraction of dirac fermions in graphene," *Nature Physics* **11**, 925–929 (2015).

[209] T. Scaffidi, N. Nandi, B. Schmidt, et al., "Hydrodynamic electron flow and hall viscosity," *Physical Review Letters* **118**, 226601 (2017).

[210] D. A. Bandurin, I. Torre, R. K. Kumar, et al., "Negative local resistance caused by viscous electron backflow in graphene," *Science* **351**, 1055–1058 (2016).

[211] L. Levitov and G. Falkovich, "Electron viscosity, current vortices and negative nonlocal resistance in graphene," *Nature Physics* **12**, 672–676 (2016).

[212] H. Guo, E. Ilseven, G. Falkovich, et al., "Higher-than-ballistic conduction of viscous electron flows," *Proceedings of the National Academy of Sciences of the United States of America* **114**, 3068–3073 (2017).

[213] L. Fritz, J. Schmalian, M. Müller, et al., "Quantum critical transport in clean graphene," *Physical Review B* **78**, 085416 (2008).

[214] L. V. Delacretaz and A. Gromov, "Transport signatures of Hall viscosity," *Physical Review Letters* **119**, 226602 (2017).

[215] F. M. D. Pellegrino, I. Torre, and M. Polini, "Nonlocal transport and the hall viscosity of two-dimensional hydrodynamic electron liquids," *Physical Review B* **96**, 195401 (2017).

[216] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, et al., "Magnetoresistance in two-component systems," *Physical Review Letters* **114**, 156601 (2015).

[217] G. Y. Vasileva, D. Smirnov, Y. L. Ivanov, et al., "Linear magnetoresistance in compensated graphene bilayer," *Physical Review B* **93**, 195430 (2016).

[218] D. K. Ki and A. F. Morpurgo, "High-quality multiterminal suspended graphene devices," *Nano Letters* **13**, 5165–5170 (2013).

[219] N. Tombros, A. Veligura, J. Junesch, et al., "Large yield production of high mobility freely suspended graphene electronic devices on a polydimethylglutarimide based organic polymer," *Journal of Applied Physics* **109**, 093702 (2011).

[220] F. Bonaccorso, A. Lombardo, T. Hasan, et al., "Production and processing of graphene and 2d crystals," *Materials Today* **15**, 564–589 (2012).

[221] A. C. Ferrari and e. al., "Science and technology roadmap for graphene, related two-dimensional crystals, and hybrid systems," *Nanoscale* **7**, 4598–4810 (2015).

[222] J. Cai, P. Ruffieux, R. Jaafar, et al., "Atomically precise bottom-up fabrication of graphene nanoribbons," *Nature* **466**, 470–473 (2010).

[223] A. Narita, X. Y. Wang, X. Feng, et al., "New advances in nanographene chemistry," *Chemical Society Reviews* **44**, 6616–6643 (2015).

[224] S. Park and R. S. Ruoff, "Chemical methods for the production of graphenes," *Nature Nanotechnology* **4**, 217–224 (2009).

[225] K. V. Emtsev, A. Bostwick, K. Horn, et al., "Towards wafer-size graphene layers by atmospheric pressure graphitization of silicon carbide," *Nature Materials* **8**, 203–207 (2009).

[226] G. Yazdi, T. Iakimov, and R. Yakimova, "Epitaxial graphene on sic: A review of growth and characterization," *Crystals* **6**, 53 (2016).

[227] C. Riedl, C. Coletti, T. Iwasaki, et al., "Quasi-free-standing epitaxial graphene on sic obtained by hydrogen intercalation," *Physical Review Letters* **103**, 246804 (2009).

[228] F. Lafont, R. Ribeiro-Palau, D. Kazazis, et al., "Quantum hall resistance standards from graphene grown by chemical vapour deposition on silicon carbide," *Nature Communications* **6**, 6806 (2015).

[229] F. Zhang, X. Chen, C. Yu, et al., "High mobility and large domain decoupled epitaxial graphene on sic (0001) surface obtained by nearly balanced hydrogen etching," *Materials Letters* **195**, 82–85 (2017).

[230] J. Yu, J. Li, W. Zhang, et al., "Synthesis of high quality two-dimensional materials via chemical vapor deposition," *Chemical Science* **6**, 6705–6716 (2015).

[231] W. Choi, N. Choudhary, G. H. Han, et al., "Recent development of two-dimensional transition metal dichalcogenides and their applications," *Materials Today* **20**, 116–130 (2017).

[232] J. H. Lee, E. K. Lee, W. J. Joo, et al., "Wafer-scale growth of single-crystal monolayer graphene on reusable hydrogen-terminated germanium," *Science* **344**, 286–289 (2014).

[233] T. Wu, X. Zhang, Q. Yuan, et al., "Fast growth of inch-sized single-crystalline graphene from a controlled single nucleus on cu-ni alloys," *Nature Materials* **15**, 43–47 (2016).

[234] Z. Yan, J. Lin, Z. Peng, et al., "Toward the synthesis of wafer-scale single-crystal graphene on copper foils," *ACS Nano* **6**, 9110–9117 (2012).

[235] L. Gan and Z. Luo, "Turning off hydrogen to realize seeded growth of subcentimeter single-crystal graphene grains on copper," *ACS Nano* **7**, 9480–9488 (2013).

[236] L. Banszerus, M. Schmitz, S. Engels, et al., "Ultrahigh-mobility graphene devices from chemical vapor deposition on reusable copper," *Science Advances* **1**, 1500222 (2015).

[237] Z. Li, Y. Wang, A. Kozbial, et al., "Effect of airborne contaminants on the wettability of supported graphene and graphite," *Nature Materials* **12**, 925–931 (2013).

[238] L. A. Belyaeva, P. M. G. van Deursen, K. I. Barbetsea, et al., "Hydrophilicity of graphene in water through transparency to polar and dispersive interactions," *Advanced Materials* **30**, 1703274 (2017).

[239] R. Wang, S. Wang, D. Zhang, et al., "Control of carrier type and density in exfoliated graphene by interface engineering," *ACS Nano* **5**, 408–412 (2011).

[240] M. Lafkioti, B. Krauss, T. Lohmann, et al., "Graphene on a hydrophobic substrate: doping reduction and hysteresis suppression under ambient conditions," *Nano Letters* **10**, 1149–1153 (2010).

[241] S. Y. Chen, P. H. Ho, R. J. Shiue, et al., "Transport/magnetotransport of high-performance graphene transistors on organic molecule-functionalized substrates," *Nano Letters* **12**, 964–969 (2012).

[242] H. Xu, Z. Zhang, R. Shi, et al., "Batch-fabricated high-performance graphene hall elements," *Sci Rep* **3**, 1207 (2013).

[243] Y. Huang, E. Sutter, N. N. Shi, et al., "Reliable exfoliation of large-area high-quality flakes of graphene and other two-dimensional materials," *ACS Nano* **9**, 10612–10620 (2015).

[244] F. Pizzocchero, L. Gammelgaard, B. S. Jessen, et al., "The hot pick-up technique for batch assembly of van der waals heterostructures," *Nature Communications* **7**, 11894 (2016).

[245] Y. Chen, X. L. Gong, and J. G. Gai, "Progress and challenges in transfer of large-area graphene films," *Advanced Science* **3**, 1500343 (2016).

[246] X. Li, W. Cai, J. An, et al., "Large-area synthesis of high-quality and uniform graphene films on copper foils," *Science* **324**, 1312–1314 (2009).

[247] S. Wang, H. Hibino, S. Suzuki, et al., "Atmospheric pressure chemical vapor deposition growth of millimeter-scale single-crystalline graphene on the copper surface with a native oxide layer," *Chemistry of Materials* **28**, 4893–4900 (2016).

[248] L. Brown, E. B. Lochocki, J. Avila, et al., "Polycrystalline graphene with single crystalline electronic structure," *Nano Letters* **14**, 5706–5711 (2014).

[249] Y. Wang, Y. Zheng, X. Xu, et al., "Electrochemical delamination of cvd-grown graphene film: toward the recyclable use of copper catalyst," *ACS Nano* **5**, 9927–9933 (2011).

[250] C. T. Cherian, F. Giustiniano, I. Martin-Fernandez, et al., "'bubble-free' electrochemical delamination of cvd graphene films," *Small* **11**, 189–194 (2015).

[251] Y. C. Lin, C. C. Lu, C. H. Yeh, et al., "Graphene annealing: how clean can it be?," *Nano Letters* **12**, 414–419 (2012).

[252] G. Algara-Siller, O. Lehtinen, A. Turchanin, et al., "Dry-cleaning of graphene," *Applied Physics Letters* **104**, 153115 (2014).

[253] D. Lee, G. Ahn, and S. Ryu, "Two-dimensional water diffusion at a graphene-silica interface," *Journal of the American Chemical Society* **136**, 6634–6642 (2014).

[254] M. Her, R. Beams, and L. Novotny, "Graphene transfer with reduced residue," *Physics Letters A* **377**, 1455–1458 (2013).

[255] J. Moser, A. Barreiro, and A. Bachtold, "Current-induced cleaning of graphene," *Applied Physics Letters* **91**, 163513 (2007).

[256] H. Wang, X. Zhang, and H. Takamatsu, "Ultraclean suspended monolayer graphene achieved by in situ current annealing," *Nanotechnology* **28**, 045706 (2017).

[257] G. Cunge, D. Ferrah, C. Petit-Etienne, et al., "Dry efficient cleaning of poly-methyl-methacrylate residues from graphene with high-density h2 and h2-n2 plasmas," *Journal of Applied Physics* **118**, 123302 (2015).

[258] A. M. Goossens, V. E. Calado, A. Barreiro, et al., "Mechanical cleaning of graphene," *Applied Physics Letters* **100**, 073110 (2012).

[259] N. Lindvall, A. Kalabukhov, and A. Yurgens, "Cleaning graphene using atomic force microscope," *Journal of Applied Physics* **111**, 064904 (2012).

[260] W. Choi, M. A. Shehzad, S. Park, et al., "Influence of removing pmma residues on surface of cvd graphene using a contact-mode atomic force microscope," *Royal Society of Chemistry Advances* **7**, 6943–6949 (2017).

[261] J.-N. Longchamp, C. Escher, and H.-W. Fink, "Ultraclean freestanding graphene by platinum-metal catalysis," *Journal of Vacuum Science and Technology B* **31**, 020605 (2013).

[262] Y. Ahn, H. Kim, Y.-H. Kim, et al., "Procedure of removing polymer residues and its influences on electronic and structural characteristics of graphene," *Applied Physics Letters* **102**, 091602 (2013).

[263] Z. H. Ni, H. M. Wang, Z. Q. Luo, et al., "The effect of vacuum annealing on graphene," *Journal of Raman Spectroscopy* **41**, 479–483 (2009).

[264] J. Hong, M. K. Park, E. J. Lee, et al., "Origin of new broad raman d and g peaks in annealed graphene," *Scientific Reports* **3**, 2700 (2013).

[265] J. E. Lee, G. Ahn, J. Shim, et al., "Optical separation of mechanical strain from charge doping in graphene," *Nature Communications* **3**, 1024 (2012).

[266] Z. Zhang, J. Du, D. Zhang, et al., "Rosin-enabled ultraclean and damage-free transfer of graphene for large-area flexible organic light-emitting diodes," *Nature Communications* **8**, 14560 (2017).

[267] J. D. Wood, G. P. Doidge, E. A. Carrion, et al., "Annealing free, clean graphene transfer using alternative polymer scaffolds," *Nanotechnology* **26**, 055302 (2015).

[268] H. Van Ngoc, Y. Qian, S. K. Han, et al., "Pmma-etching-free transfer of wafer-scale chemical vapor deposition two-dimensional atomic crystal by a water soluble polyvinyl alcohol polymer method," *Scientific Reports* **6**, 33096 (2016).

[269] H. H. Kim, B. Kang, J. W. Suk, et al., "Clean transfer of wafer-scale graphene via liquid phase removal of polycyclic aromatic hydrocarbons," *ACS Nano* **9**, 4726–4733 (2015).

[270] W. H. Lin, T. H. Chen, J. K. Chang, et al., "A direct and polymer-free method for transferring graphene grown by chemical vapor deposition to any substrate," *ACS Nano* **8**, 1784–1791 (2014).

[271] B. Wang, M. Huang, L. Tao, et al., "Support-free transfer of ultrasmooth graphene films facilitated by self-assembled monolayers for electronic devices and patterns," *ACS Nano* **10**, 1404–1410 (2016).

[272] L. Huang, H. Xu, Z. Zhang, et al., "Graphene/si cmos hybrid hall integrated circuits," *Sci Rep* **4**, 5548 (2014).

[273] K. S. Novoselov, A. Mishchenko, A. Carvalho, et al., "2d materials and van der waals heterostructures," *Science* **353**, 9439 (2016).

[274] Y. Liu, N. O. Weiss, X. Duan, et al., "Van der waals heterostructures and devices," *Nature Reviews Materials* **1**, 16042 (2016).

[275] C. P. Lu, G. Li, K. Watanabe, et al., "Mos 2 mos2: choice substrate for accessing and tuning the electronic properties of graphene," *Physical Review Letters* **113**, 156804 (2014).

[276] G. Cassabois, P. Valvin, and B. Gil, "Hexagonal boron nitride is an indirect bandgap semiconductor," *Nature Photonics* **10**, 262–266 (2016).

[277] Y. Hattori, T. Taniguchi, K. Watanabe, et al., "Layer-by-layer dielectric breakdown of hexagonal boron nitride," *ACS Nano* **9**, 916–921 (2015).

[278] Y. Hattori, T. Taniguchi, K. Watanabe, et al., "Anisotropic dielectric breakdown strength of single crystal hexagonal boron nitride," *ACS Applied Materials Interfaces* **8**, 27877–27884 (2016).

[279] T. Zhang, M.-Q. Wu, S.-R. Zhang, et al., "Permittivity and its temperature dependence in hexagonal structure bn dominated by the local electric field," *Chinese Physics B* **21**, 077701 (2012).

[280] R. V. Gorbachev, I. Riaz, R. R. Nair, et al., "Hunting for monolayer boron nitride: optical and raman signatures," *Small* **7**, 465–468 (2011).

[281] T. Taniguchi and K. Watanabe, "Synthesis of high-purity boron nitride single crystals under high pressure by using ba–bn solvent," *Journal of Crystal Growth* **303**, 525–529 (2007).

[282] K. Watanabe, T. Taniguchi, and H. Kanda, "Direct-bandgap properties and evidence for ultraviolet lasing of hexagonal boron nitride single crystal," *Nature Materials* **3**, 404–409 (2004).

[283] M. H. Khan, H. K. Liu, X. Sun, et al., "Few-atomic-layered hexagonal boron nitride: Cvd growth, characterization, and applications," *Materials Today* **20**, 611–628 (2017).

[284] W. Gannett, W. Regan, K. Watanabe, et al., "Boron nitride substrates for high mobility chemical vapor deposited graphene," *Applied Physics Letters* **98**, 242105 (2011).

[285] N. Petrone, C. R. Dean, I. Meric, et al., "Chemical vapor deposition-derived graphene with electrical performance of exfoliated graphene," *Nano Letters* **12**, 2751–2756 (2012).

[286] X. Song, J. Gao, Y. Nie, et al., "Chemical vapor deposition growth of large-scale hexagonal boron nitride with controllable orientation," *Nano Research* **8**, 3164–3176 (2015).

[287] K. K. Kim, A. Hsu, X. Jia, et al., "Synthesis of monolayer hexagonal boron nitride on cu foil using chemical vapor deposition," *Nano Letters* **12**, 161–166 (2012).

[288] G. Lu, T. Wu, Q. Yuan, et al., "Synthesis of large single-crystal hexagonal boron nitride grains on cu-ni alloy," *Nature Communications* **6**, 6160 (2015).

[289] S. Caneva, R. S. Weatherup, B. C. Bayer, et al., "Nucleation control for large, single crystalline domains of monolayer hexagonal boron nitride via si-doped fe catalysts," *Nano Letters* **15**, 1867–1875 (2015).

[290] J. H. Park, J. C. Park, S. J. Yun, et al., "Large-area monolayer hexagonal boron nitride on pt foil," *ACS Nano* **8**, 8520–8528 (2014).

[291] Y. D. Kim, M. H. Bae, J. T. Seo, et al., "Focused-laser-enabled p-n junctions in graphene field-effect transistors," *ACS Nano* **7**, 5850–5857 (2013).

[292] K. K. Kim, A. Hsu, X. Jia, et al., "Synthesis and characterization of hexagonal boron nitride film as a dielectric layer for graphene devices," *ACS Nano* **6**, 8583–8590 (2012).

[293] A. R. Jang, S. Hong, C. Hyun, et al., "Wafer-scale and wrinkle-free epitaxial growth of single-orientated multilayer hexagonal boron nitride on sapphire," *Nano Letters* **16**, 3360–3366 (2016).

[294] Y. Shi, C. Hamsen, X. Jia, et al., "Synthesis of few-layer hexagonal boron nitride thin film by chemical vapor deposition," *Nano Letters* **10**, 4134–4139 (2010).

[295] S. M. Kim, A. Hsu, M. H. Park, et al., "Synthesis of large-area multilayer hexagonal boron nitride for high material performance," *Nature Communications* **6**, 8662 (2015).

[296] J. Ohta and H. Fujioka, "Sputter synthesis of wafer-scale hexagonal boron nitride films via interface segregation," *Applied Physics Letters Materials* **5**, 076107 (2017).

[297] S. Sonde, A. Dolocan, N. Lu, et al., "Ultrathin, wafer-scale hexagonal boron nitride on dielectric surfaces by diffusion and segregation mechanism," *2D Materials* **4**, 025052 (2017).

[298] A. Ismach, H. Chou, D. A. Ferrer, et al., "Toward the controlled synthesis of hexagonal boron nitride films," *ACS Nano* **6**, 6378–6385 (2012).

[299] R. Y. Tay, S. H. Tsang, M. Loeblein, et al., "Direct growth of nanocrystalline hexagonal boron nitride films on dielectric substrates," *Applied Physics Letters* **106**, 101901 (2015).

[300] V. Shautsova, A. M. Gilbertson, N. C. Black, et al., "Hexagonal boron nitride assisted transfer and encapsulation of large area cvd graphene," *Scientific Reports* **6**, 30210 (2016).

[301] A. Dankert, B. Karpiak, and S. P. Dash, "Hall sensors batch-fabricated on all-cvd h-bn/graphene/h-bn heterostructures," *Sci Rep* **7**, 15231 (2017).

[302] M. W. Iqbal, M. Z. Iqbal, X. Jin, et al., "Superior characteristics of graphene field effect transistor enclosed by chemical-vapor-deposition-grown hexagonal boron nitride," *Journal of Materials Chemistry C* **2**, 7776–7784 (2014).

[303] J. Y. Tan, A. Avsar, J. Balakrishnan, et al., "Electronic transport in graphene-based heterostructures," *Applied Physics Letters* **104**, 183504 (2014).

[304] R. H. Miwa, T. M. Schmidt, W. L. Scopel, et al., "Doping of graphene adsorbed on the a-sio2 surface," *Applied Physics Letters* **99**, 163108 (2011).

[305] H. Wang, T. Taychatanapat, A. Hsu, et al., "Bn/graphene/bn transistors for rf applications," *IEEE Electron Device Letters* **32**, 1209–1211 (2011).

[306] P. Avouris, T. Heinz, and T. Low, eds., *2D Materials Properties and Devices*. Cambridge University Press, 1 ed. (2017).

[307] Y. Cao, A. Mishchenko, G. L. Yu, et al., "Quality heterostructures from two-dimensional crystals unstable in air by their assembly in inert atmosphere," *Nano Letters* **15**, 4914–4921 (2015).

[308] A. Avsar, I. J. Vera-Marun, J. Y. Tan, et al., "Air-stable transport in graphene-contacted, fully encapsulated ultrathin black phosphorus-based field-effect transistors," *ACS Nano* **9**, 4138–4145 (2015).

[309] E. Khestanova, F. Guinea, L. Fumagalli, et al., "Universal shape and pressure inside bubbles appearing in van der waals heterostructures," *Nature Communications* **7**, 12587 (2016).

[310] K. S. Vasu, E. Prestat, J. Abraham, et al., "Van der waals pressure and its effect on trapped interlayer molecules," *Nature Communications* **7**, 12168 (2016).

[311] N. Leconte, H. Kim, H. J. Kim, et al., "Graphene bubbles and their role in graphene quantum transport," *Nanoscale* **9**, 6041–6047 (2017).

[312] X. Cui, G. H. Lee, Y. D. Kim, et al., "Multi-terminal transport measurements of mos2 using a van der waals heterostructure device platform," *Nature Nanotechnology* **10**, 534–540 (2015).

[313] S. Masubuchi, M. Morimoto, S. Morikawa, et al., "Autonomous robotic searching and assembly of two-dimensional crystals to build van der waals superlattices," *Nature Communications* **9**, 1413 (2018).

[314] D. Wang, G. Chen, C. Li, et al., "Thermally induced graphene rotation on hexagonal boron nitride," *Physical Review Letters* **116**, 126101 (2016).

[315] L. Wang, Z. Chen, C. R. Dean, et al., "Negligible environmental sensitivity of graphene in a hexagonal boron nitride/graphene/h-bn sandwich structure," *ACS Nano* **6**, 9314–9319 (2012).

[316] S. Engels, B. Terrés, F. Klein, et al., "Impact of thermal annealing on graphene devices encapsulated in hexagonal boron nitride," *physica status solidi b* **251**, 2545–2550 (2014).

[317] P. J. Zomer, M. H. D. Guimarães, J. C. Brant, et al., "Fast pick up technique for high quality heterostructures of bilayer graphene and hexagonal boron nitride," *Applied Physics Letters* **105**, 013101 (2014).

[318] G. H. Lee, Y. J. Yu, X. Cui, et al., "Flexible and transparent mos2 field-effect transistors on hexagonal boron nitride-graphene heterostructures," *ACS Nano* **7**, 7931–7936 (2013).

[319] P. J. Zomer, S. P. Dash, N. Tombros, et al., "A transfer technique for high mobility graphene devices on commercially available hexagonal boron nitride," *Applied Physics Letters* **99**, 232104 (2011).

[320] D. G. Purdie, N. M. Pugno, T. Taniguchi, et al., "Cleaning Interfaces in Layered Materials Heterostructures." arXiv:1803.00912 [cond-mat.mes-hall].

[321] A. R. Cadore, E. Mania, K. Watanabe, et al., "Thermally activated hysteresis in high quality graphene/h-bn devices," *Applied Physics Letters* **108**, 233101 (2016).

[322] J. Mohrmann, K. Watanabe, T. Taniguchi, et al., "Persistent hysteresis in graphene-mica van der waals heterostructures," *Nanotechnology* **26**, 015202 (2015).

[323] H. Wang, Y. Wu, C. Cong, et al., "Hysteresis of electronic transport in graphene transistors," *ACS Nano* **4**, 7221–7228 (2010).

[324] L. Ju, J. Velasco, J., E. Huang, et al., "Photoinduced doping in heterostructures of graphene and boron nitride," *Nature Nanotechnology* **9**, 348–352 (2014).

[325] C. Neumann, L. Rizzi, S. Reichardt, et al., "Spatial control of laser-induced doping profiles in graphene on hexagonal boron nitride," *ACS Applied Materials Interfaces* **8**, 9377–9383 (2016).

[326] F. Amet, A. J. Bestwick, J. R. Williams, et al., "Composite fermions and broken symmetries in graphene," *Nature Communications* **6**, 5838 (2015).

[327] P. Maher, L. Wang, Y. Gao, et al., "Bilayer graphene. tunable fractional quantum hall phases in bilayer graphene," *Science* **345**, 61–64 (2014).

[328] M. S. Dresselhaus, A. Jorio, M. Hofmann, et al., "Perspectives on carbon nanotubes and graphene raman spectroscopy," *Nano Letters* **10**, 751–758 (2010).

[329] A. C. Ferrari, J. C. Meyer, V. Scardaci, et al., "Raman spectrum of graphene and graphene layers," *Physical Review Letters* **97**, 187401 (2006).

[330] A. C. Ferrari and D. M. Basko, "Raman spectroscopy as a versatile tool for studying the properties of graphene," *Nature Nanotechnology* **8**, 235–246 (2013).

[331] L. M. Malard, M. A. Pimenta, G. Dresselhaus, et al., "Raman spectroscopy in graphene," *Physics Reports* **473**, 51–87 (2009).

[332] S. Berciaud, X. Li, H. Htoon, et al., "Intrinsic line shape of the raman 2d-mode in freestanding graphene monolayers," *Nano Letters* **13**, 3517–3523 (2013).

[333] P. H. Tan, W. P. Han, W. J. Zhao, et al., "The shear mode of multilayer graphene," *Nature Materials* **11**, 294–300 (2012).

[334] J. B. Wu, Z. X. Hu, X. Zhang, et al., "Interface coupling in twisted multilayer graphene by resonant raman spectroscopy of layer breathing modes," *ACS Nano* **9**, 7440–7449 (2015).

[335] D. H. Chae, B. Krauss, K. von Klitzing, et al., "Hot phonons in an electrically biased graphene constriction," *Nano Letters* **10**, 466–471 (2010).

[336] K. T. Nguyen, D. Abdula, C. L. Tsai, et al., "Temperature and gate voltage dependent raman spectra of single-layer graphene," *ACS Nano* **5**, 5273–5279 (2011).

[337] D. Yoon, Y. W. Son, and H. Cheong, "Negative thermal expansion coefficient of graphene measured by raman spectroscopy," *Nano Letters* **11**, 3227–3231 (2011).

[338] C. Casiraghi, S. Pisana, K. S. Novoselov, et al., "Raman fingerprint of charged impurities in graphene," *Applied Physics Letters* **91**, 233108 (2007).

[339] J. Yan, Y. Zhang, P. Kim, et al., "Electric field effect tuning of electron-phonon coupling in graphene," *Physical Review Letters* **98**, 166802 (2007).

[340] N. Ferralis, "Probing mechanical properties of graphene with raman spectroscopy," *Journal of Materials Science* **45**, 5135–5149 (2010).

[341] J. Zabel, R. R. Nair, A. Ott, et al., "Raman spectroscopy of graphene and bilayer under biaxial strain: bubbles and balloons," *Nano Letters* **12**, 617–621 (2012).

[342] K. Elibol, B. C. Bayer, S. Hummel, et al., "Visualising the strain distribution in suspended two-dimensional materials under local deformation," *Scientific Reports* **6**, 28485 (2016).

[343] G. Ahn, H. R. Kim, T. Y. Ko, et al., "Optical probing of the electronic interaction between graphene and hexagonal boron nitride," *ACS Nano* **7**, 1533–1541 (2013).

[344] A. Das, B. Chakraborty, S. Piscanec, et al., "Phonon renormalization in doped bilayer graphene," *Physical Review B* **79**, 155417 (2009).

[345] F. Forster, A. Molina-Sanchez, S. Engels, et al., "Dielectric screening of the kohn anomaly of graphene on hexagonal boron nitride," *Physical Review B* **88**, 085419 (2013).

[346] D. Yoon, H. Moon, Y.-W. Son, et al., "Interference effect on raman spectrum of graphene on sio2/si," *Physical Review B* **80**, 125422 (2009).

[347] E. Ramsden, *Hall-Effect Sensors*. Newnes, 2 ed. (2006).

[348] L. Huang, Z. Zhang, B. Chen, et al., "Ultra-sensitive graphene hall elements," *Applied Physics Letters* **104**, 183106 (2014).

[349] J. Dauber, A. A. Sagade, M. Oellers, et al., "Ultra-sensitive hall sensors based on graphene encapsulated in hexagonal boron nitride," *Applied Physics Letters* **106**, 193501 (2015).

[350] A. T. Neal, H. Liu, J. Gu, et al., "Magneto-transport in mos2: phase coherence, spin-orbit scattering, and the hall factor," *ACS Nano* **7**, 7077–7082 (2013).

[351] D. B. Farmer, V. Perebeinos, Y.-M. Lin, et al., "Charge trapping and scattering in epitaxial graphene," *Physical Review B* **84**, 205417 (2011).

[352] H. K. Tang, J. N. Leaw, J. N. B. Rodrigues, et al., "The role of electron-electron interactions in two-dimensional dirac fermions," *Science* **361**, 570–574 (2018).

[353] A. S. Mayorov, D. C. Elias, M. Mucha-Kruczynski, et al., "Interaction-driven spectrum reconstruction in bilayer graphene," *Science* **333**, 860–863 (2011).

[354] H. Rokni and W. Lu, "Layer-by-layer insight into electrostatic charge distribution of few-layer graphene," *Scientific Reports* **7**, 42821 (2017).

[355] C. Cobaleda, S. Pezzini, E. Diez, et al., "Temperature- and density-dependent transport regimes in ah-bn/bilayer graphene/h-bn heterostructure," *Physical Review B* **89**, 121404 (2014).

[356] J. Li, L. Z. Tan, K. Zou, et al., "Effective mass in bilayer graphene at low carrier densities: The role of potential disorder and electron-electron interaction," *Physical Review B* **94**, 161406 (2016).

[357] F. Chiappini, S. Wiedmann, M. Titov, et al., "Magnetotransport in single-layer graphene in a large parallel magnetic field," *Physical Review B* **94**, 085302 (2016).

[358] F. Kisslinger, C. Ott, C. Heide, et al., "Linear magnetoresistance in mosaic-like bilayer graphene," *Nature Physics* **11**, 650–653 (2015).

[359] P. S. Alekseev, A. P. Dmitriev, I. V. Gornyi, et al., "Magnetoresistance of compensated semimetals in confined geometries," *Physical Review B* **95**, 165410 (2017).

[360] J. Lu, H. Zhang, W. Shi, et al., "Graphene magnetoresistance device in van der pauw geometry," *Nano Letters* **11**, 2973–2977 (2011).

[361] B. Zhou, K. Watanabe, T. Taniguchi, et al., "Extraordinary magnetoresistance in encapsulated monolayer graphene devices," *Applied Physics Letters* **116**, 053102 (2020).

[362] K. Gopinadhan, Y. J. Shin, R. Jalil, et al., "Extremely large magnetoresistance in few-layer graphene/boron-nitride heterostructures," *Nature Communications* **6**, 8337 (2015).

[363] P. S. Alekseev, "Negative magnetoresistance in viscous flow of two-dimensional electrons," *Physical Review Letters* **117**, 166601 (2016).

[364] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, et al., "Measuring Hall Viscosity of Graphene's Electron Fluid." arXiv:1806.01606 [cond-mat.mes-hall].

[365] A. Dmitriev, M. Dyakonov, and R. Jullien, "Classical mechanism for negative magnetoresistance in two dimensions," *Physical Review B* **64**, 233321 (2001).

[366] A. D. Mirlin, D. G. Polyakov, F. Evers, et al., "Quasiclassical negative magnetoresistance of a 2d electron gas: interplay of strong scatterers and smooth disorder," *Physical Review Letters* **87**, 126805 (2001).

[367] A. Grosz, *High Sensitivity Magnetometers*. Springer Nature, 1 ed. (2017).

[368] S. A. Solin, T. Thio, D. R. Hines, et al., "Enhanced room-temperature geometric magnetoresistance in inhomogeneous narrow-gap semiconductors," *Science* **289**, 1530–1532 (2000).

[369] L. Jogschies, D. Klaas, R. Kruppe, et al., "Recent developments of magnetoresistive sensors for industrial applications," *Sensors* **15**, 28665–28689 (2015).

[370] P. Ripka and M. Janosek, "Advances in magnetic field sensors," *IEEE Sensors Journal* **10**, 1108–1116 (2010).

[371] M. N. Baibich, J. M. Broto, A. Fert, et al., "Giant magnetoresistance of (001)fe/(001)cr magnetic superlattices," *Physical Review Letters* **61**, 2472–2475 (1988).

[372] G. Binasch, P. Grunberg, F. Saurenbach, et al., "Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange," *Physical Review B* **39**, 4828–4830 (1989).

[373] J.-G. Zhu and C. Park, "Magnetic tunnel junctions," *Materials Today* **9**, 36–45 (2006).

[374] S. Yuasa, T. Nagahama, A. Fukushima, et al., "Giant room-temperature magnetoresistance in single-crystal fe/mgo/fe magnetic tunnel junctions," *Nature Materials* **3**, 868–871 (2004).

[375] S. S. Parkin, C. Kaiser, A. Panchula, et al., "Giant tunnelling magnetoresistance at room temperature with mgo (100) tunnel barriers," *Nature Materials* **3**, 862–867 (2004).

[376] J. Paul, C. Schnieders, J. Traute, et al., "Sensors based on tunnel magnetoresistance - new technology, new opportunities," *AMA Sensor 2015 Proceedings* **2015**, 234–239 (2015).

[377] S. Pisana, P. M. Braganca, E. E. Marinero, et al., "Tunable nanoscale graphene magnetometers," *Nano Letters* **10**, 341–346 (2010).

[378] G. Song, M. Ranjbar, and R. A. Kiehl, "Operation of graphene magnetic field sensors near the charge neutrality point," *Communications Physics* **2** (2019).

[379] R. S. Popovic, "High resolution hall magnetic sensors," *29th International Conference on Microelectronics Proceedings*69–74 (2014).

[380] S. Sanfilippo, "Hall probes: physics and application to magnetometry." arXiv:1103.1271v1 [physics.acc-ph].

[381] P. Daniil and E. Cohen, "Low field hall effect magnetometry," *Journal of Applied Physics* **53**, 8257–8259 (1982).

[382] P. J. A. Munter, "A low-offset spinning-current hall plate," *Sensors and Actuators A: Physical* **22**, 743–746 (1990).

[383] A. Kerlain and V. Mosser, "Low frequency noise suppression in iii-v hall magnetic microsystems with integrated switches," *Sensor Letters* **5**, 192–195 (2007).

[384] W. Chen, *The Electrical Engineering Handbook*. Elsevier, 1 ed. (2005).

[385] V. Mosser, N. Matringe, and Y. Haddab, "A spinning current circuit for hall measurements down to the nanotesla range," *IEEE Transactions on Instrumentation and Measurement* **66**, 637–650 (2017).

[386] A. Kaverzin, *Electronic transport and flicker noise in graphene structures.* PhD thesis, University of Exeter (2011).

[387] F. N. Hooge, "1/f noise sources," *IEEE Transactions on Electron Devices* **41**, 1926–1935 (1994).

[388] X.-y. Chen, *Lattice Scattering and 1/f Noise in Semiconductors.* PhD thesis, Technical University Eindhoven (1997).

[389] A. A. Balandin, "Low-frequency 1/f noise in graphene devices," *Nature Nanotechnology* **8**, 549–555 (2013).

[390] F. N. Hooge, "1/f noise," *Physica B+C* **83**, 14–23 (1976).

[391] P. Dutta and P. M. Horn, "Low-frequency fluctuations in solids:1fnoise," *Reviews of Modern Physics* **53**, 497–516 (1981).

[392] P. Karnatak, T. Paul, S. Islam, et al., "1/f noise in van der waals materials and hybrids," *Advances in Physics: X* **2**, 428–449 (2017).

[393] L. Ren and J. S. Liberis, "1/f-noise in the hall voltage of epitaxial n-gaas," *Physica B: Condensed Matter* **183**, 40–44 (1993).

[394] T. G. M. Kleinpenning, "1/f noise in hall effect: Fluctuations in mobility," *Journal of Applied Physics* **51**, 3438–3438 (1980).

[395] J. Lu, J. Pan, S.-S. Yeh, et al., "Negative correlation between charge carrier density and mobility fluctuations in graphene," *Physical Review B* **90**, 085434 (2014).

[396] H. M. J. Vaes and T. G. M. Kleinpenning, "Hall-effect noise in semiconductors," *Journal of Applied Physics* **48**, 5131–5134 (1977).

[397] B. T. Schaefer, L. Wang, A. Jarjour, et al., "Magnetic field detection limits for ultraclean graphene hall sensors," *Nature Communications* **11**, 4163 (2020).

[398] T. Ciuk, O. Petruk, A. Kowalik, et al., "Low-noise epitaxial graphene on sic hall effect element for commercial applications," *Applied Physics Letters* **108**, 223504 (2016).

[399] H. Xu, L. Huang, Z. Zhang, et al., "Flicker noise and magnetic resolution of graphene hall sensors at low frequency," *Applied Physics Letters* **103**, 112405 (2013).

[400] P. Tian, *Graphene-Based High Spatial Resolution Hall Sensors with Potential Application for Data Storage Media Characterisation.* PhD thesis, University of Manchester (2014).

[401] C. Kumar, M. Kuiri, J. Jung, et al., "Tunability of 1/f noise at multiple dirac cones in hbn encapsulated graphene devices," *Nano Letters* **16**, 1042–1049 (2016).

[402] S. Takeshita, S. Matsuo, T. Tanaka, et al., "Anomalous behavior of 1/f noise in graphene near the charge neutrality point," *Applied Physics Letters* **108**, 103106 (2016).

[403] G. Xu, J. Torres, C. M., Y. Zhang, et al., "Effect of spatial charge inhomogeneity on 1/f noise behavior in graphene," *Nano Letters* **10**, 3312–3317 (2010).

[404] Y. Zhang, E. E. Mendez, and X. Du, "Mobility-dependent low-frequency noise in graphene field-effect transistors," *ACS Nano* **5**, 8124–8130 (2011).

[405] A. Rahman, J. W. Guikema, and N. Markovic, "Quantum noise and asymmetric scattering of electrons and holes in graphene," *Nano Letters* **14**, 6621–6625 (2014).

[406] M. Kayyalha and Y. P. Chen, "Observation of reduced 1/f noise in graphene field effect transistors on boron nitride substrates," *Applied Physics Letters* **107**, 113101 (2015).

[407] M. A. Stolyarov, G. Liu, S. L. Rumyantsev, et al., "Suppression of 1/f noise in near-ballistic h-bn-graphene-h-bn heterostructure field-effect transistors," *Applied Physics Letters* **107**, 023106 (2015).

[408] Y. Wang, V. X. Ho, Z. N. Henschel, et al., "Effect of high-k dielectric layer on 1/f noise behavior in graphene field-effect transistors," *ACS Applied Nano Materials* **4**, 3647–3653 (2021).

[409] S. Kakkar, P. Karnatak, M. Ali Aamir, et al., "Optimal architecture for ultralow noise graphene transistors at room temperature," *Nanoscale* **12**, 17762–17768 (2020).

[410] G. Liu, S. Rumyantsev, M. S. Shur, et al., "Origin of 1/f noise in graphene multilayers: Surface vs. volume," *Applied Physics Letters* **102**, 093111 (2013).

[411] A. N. Pal, S. Ghatak, V. Kochat, et al., "Microscopic mechanism of 1/f noise in graphene: role of energy band dispersion," *ACS Nano* **5**, 2075–2081 (2011).

[412] M. Kumar, A. Laitinen, D. Cox, et al., "Ultra low 1/f noise in suspended bilayer graphene," *Applied Physics Letters* **106**, 263505 (2015).

[413] P. Karnatak, T. P. Sai, S. Goswami, et al., "Current crowding mediated large contact noise in graphene field-effect transistors," *Nature Communications* **7**, 13703 (2016).

[414] A. Allain, J. Kang, K. Banerjee, et al., "Electrical contacts to two-dimensional semiconductors," *Nature Materials* **14**, 1195–205 (2015).

[415] T. Chari, R. Ribeiro-Palau, C. R. Dean, et al., "Resistivity of rotated graphite-graphene contacts," *Nano Letters* **16**, 4477–4482 (2016).

[416] A. A. Kaverzin, A. S. Mayorov, A. Shytov, et al., "Impurities as a source of1/fnoise in graphene," *Physical Review B* **85**, 075435 (2012).

[417] A. K. Behera, C. T. Harris, D. V. Pete, et al., "High-performance and ultralow-noise two-dimensional heterostructure field-effect transistors with one-dimensional electrical contacts," *ACS Applied Electronic Materials* **3**, 4126–4134 (2021).

[418] H. He, N. Shetty, T. Bauch, et al., "The performance limits of epigraphene hall sensors doped across the dirac point," *Applied Physics Letters* **116**, 223504 (2020).

[419] G. Liu, W. Stillman, S. Rumyantsev, et al., "Low-frequency electronic noise in the double-gate single-layer graphene transistors," *Applied Physics Letters* **95**, 033103 (2009).

[420] A. Balandin, K. L. Wang, A. Svizhenko, et al., "The fundamental 1/f noise and the hooge parameter in semiconductor quantum wires," *IEEE Transactions on Electron Devices* **46**, 1240–1244 (1999).

[421] H. C. de Graaff and M. T. M. Huybers, "1/fnoise in polycrystalline silicon resistors," *Journal of Applied Physics* **54**, 2504–2507 (1983).

[422] V. Palenskis and K. Maknys, "Nature of low-frequency noise in homogeneous semiconductors," *Scientific Reports* **5**, 18305 (2015).

[423] J. Dobbert, L. Tran, F. Hatami, et al., "Low frequency noise in insb/gaas and insb/si channels," *Applied Physics Letters* **97**, 102101 (2010).

[424] J. Dobbert, L. Tran, F. Hatami, et al., "A comparison of the low frequency noise in insb grown on gaas and si by mbe," *Journal of Crystal Growth* **323**, 393–396 (2011).

[425] G. Liu, S. Rumyantsev, M. Shur, et al., "Graphene thickness-graded transistors with reduced electronic noise," *Applied Physics Letters* **100**, 033103 (2012).

[426] C.-C. Tang, M.-Y. Li, L. J. Li, et al., "Characteristics of a sensitive micro-hall probe fabricated on chemical vapor deposited graphene over the temperature range from liquid-helium to room temperature," *Applied Physics Letters* **99**, 112107 (2011).

[427] S. Sonusen, O. Karci, M. Dede, et al., "Single layer graphene hall sensors for scanning hall probe microscopy (shpm) in 3–300k temperature range," *Applied Surface Science* **308**, 414–418 (2014).

[428] C. Cho, S. K. Lee, J. W. Noh, et al., "Contact resistance improvement by the modulation of peripheral length to area ratio of graphene contact pattern," *Applied Physics Letters* **106**, 213107 (2015).

[429] F. Xia, V. Perebeinos, Y. M. Lin, et al., "The origins and limits of metal-graphene junction resistance," *Nature Nanotechnology* **6**, 179–184 (2011).

[430] D. Collomb, P. Li, and S. J. Bending, "Nanoscale graphene hall sensors for high-resolution ambient magnetic imaging," *Sci Rep* **9**, 14424 (2019).

[431] V. Panchal, K. Cedergren, R. Yakimova, et al., "Small epitaxial graphene devices for magnetosensing applications," *Journal of Applied Physics* **111**, 07E509 (2012).

[432] G. Boero, M. Demierre, P. A. Besse, et al., "Micro-hall devices: performance, technologies and applications," *Sensors and Actuators A: Physical* **106**, 314–320 (2003).

[433] K. Vervaeke, E. Simoen, G. Borghs, et al., "Size dependence of microscopic hall sensor detection limits," *Rev Sci Instrum* **80**, 074701 (2009).

[434] M. Bando, T. Ohashi, M. Dede, et al., "High sensitivity and multifunctional micro-hall sensors fabricated using inalsb/inassb/inalsb heterostructures," *Journal of Applied Physics* **105**, 07E909 (2009).

[435] V. P. Kunets, W. T. Black, Y. I. Mazur, et al., "Highly sensitive micro-hall devices based on alinsb insb heterostructures," *Journal of Applied Physics* **98**, 014506 (2005).

[436] J. Hwang, M. Kim, D. Campbell, et al., "van der waals epitaxial growth of graphene on sapphire by chemical vapor deposition without a metal catalyst," *ACS Nano* **7**, 385–395 (2013).

[437] J. Sun, T. Gao, X. Song, et al., "Direct growth of high-quality graphene on high-kappa dielectric srtio(3) substrates," *Journal American Chemical Society* **136**, 6574–6577 (2014).

[438] J. Mischke, J. Pennings, E. Weisenseel, et al., "Direct growth of graphene on gan via plasma-enhanced chemical vapor deposition under n2 atmosphere," *2D Materials* **7**, 035019 (2020).

[439] L. Gao, W. Ren, H. Xu, et al., "Repeated growth and bubbling transfer of graphene with millimetre-size single-crystal grains using platinum," *Nature Communications* **3**, 699 (2012).

[440] T. Kobayashi, M. Bando, N. Kimura, et al., "Production of a 100-m-long high-quality graphene transparent conductive film by roll-to-roll chemical vapor deposition and transfer process," *Applied Physics Letters* **102**, 023112 (2013).

[441] S. Bae, H. Kim, Y. Lee, et al., "Roll-to-roll production of 30-inch graphene films for transparent electrodes," *Nature Nanotechnology* **5**, 574–578 (2010).

[442] S. Rahimi, L. Tao, S. F. Chowdhury, et al., "Toward 300 mm wafer-scalable high-performance polycrystalline chemical vapor deposited graphene transistors," *ACS Nano* **8**, 10471–10479 (2014).

[443] J. Lee, Y. Kim, H. J. Shin, et al., "Crack-release transfer method of wafer-scale grown graphene onto large-area substrates," *ACS Applied Material Interfaces* **6**, 12588–12593 (2014).

[444] A. J. Watson, W. Lu, M. H. D. Guimarães, et al., "Transfer of large-scale two-dimensional semiconductors: challenges and developments," *2D Materials* **8**, 032001 (2021).

[445] J. Sonntag, J. Li, A. Plaud, et al., "Excellent electronic transport in heterostructures of graphene and monoisotopic boron-nitride grown at atmospheric pressure," *2D Materials* **7**, 031009 (2020).

[446] J. Sun, C. Lu, Y. Song, et al., "Recent progress in the tailored growth of two-dimensional hexagonal boron nitride via chemical vapour deposition," *Chemical Society Review* **47**, 4242–4257 (2018).

[447] K. K. Kim, H. S. Lee, and Y. H. Lee, "Synthesis of hexagonal boron nitride heterostructures for 2d van der waals electronics," *Chemical Society Review* **47**, 6342–6369 (2018).

[448] H. Wang, F. Liu, W. Fu, et al., "Two-dimensional heterostructures: fabrication, characterization, and application," *Nanoscale* **6**, 12250–12272 (2014).

[449] A. Quellmalz, X. Wang, S. Sawallich, et al., "Large-area integration of two-dimensional materials and their heterostructures by wafer bonding," *Nature Communications* **12**, 917 (2021).

[450] W. Yang, G. Chen, Z. Shi, et al., "Epitaxial growth of single-domain graphene on hexagonal boron nitride," *Nature Materials* **12**, 792–797 (2013).

[451] C. Zhang, S. Zhao, C. Jin, et al., "Direct growth of large-area graphene and boron nitride heterostructures by a co-segregation method," *Nature Communications* **6**, 6519 (2015).

[452] M. Wang, M. Kim, D. Odkhuu, et al., "Catalytic transparency of hexagonal boron nitride on copper for chemical vapor deposition growth of large-area and high-quality graphene," *ACS Nano* **8**, 5478–5483 (2014).

[453] K. Kim, J. Y. Choi, T. Kim, et al., "A role for graphene in silicon-based semiconductor devices," *Nature* **479**, 338–344 (2011).

[454] Y. M. Lin, A. Valdes-Garcia, S. J. Han, et al., "Wafer-scale graphene integrated circuit," *Science* **332**, 1294–1297 (2011).

[455] D. Neumaier, S. Pindl, and M. C. Lemme, "Integrating graphene into semiconductor fabrication lines," *Nature Materials* **18**, 525–529 (2019).

[456] S. Wittmann, C. Glacer, S. Wagner, et al., "Graphene membranes for hall sensors and microphones integrated with cmos-compatible processes," *ACS Applied Nano Materials* **2**, 5079–5085 (2019).

[457] A. A. Sagade, D. Neumaier, D. Schall, et al., "Highly air stable passivation of graphene based field effect devices," *Nanoscale* **7**, 3558–3564 (2015).

[458] G. Lupina, J. Kitzmann, I. Costina, et al., "Residual metallic contamination of transferred chemical vapor deposited graphene," *ACS Nano* **9**, 4776–4785 (2015).

[459] S. Goossens, G. Navickaite, C. Monasterio, et al., "Broadband image sensor array based on graphene–cmos integration," *Nature Photonics* **11**, 366–371 (2017).

[460] F. Xia, H. Wang, D. Xiao, et al., "Two-dimensional material nanophotonics," *Nature Photonics* **8**, 899–907 (2014).

[461] A. F. Morpurgo and F. Guinea, "Intervalley scattering, long-range disorder, and effective time-reversal symmetry breaking in graphene," *Physical Review Letters* **97**, 196804 (2006).

[462] F. V. Tikhonenko, D. W. Horsell, R. V. Gorbachev, et al., "Weak localization in graphene flakes," *Physical Review Letters* **100**, 056802 (2008).

[463] K. Kechedzhi, V. I. Fal'ko, E. McCann, et al., "Influence of trigonal warping on interference effects in bilayer graphene," *Physical Review Letters* **98**, 176806 (2007).

[464] B. Jouault, B. Jabakhanji, N. Camara, et al., "Interplay between interferences and electron-electron interactions in epitaxial graphene," *Physical Review B* **83**, 195417 (2011).

[465] K. S. Novoselov, E. McCann, S. V. Morozov, et al., "Unconventional quantum hall effect and berry's phase of 2pi in bilayer graphene," *Nature Physics* **2**, 177–180 (2006).

[466] K. S. Novoselov, Z. Jiang, Y. Zhang, et al., "Room-temperature quantum hall effect in graphene," *Science* **315**, 1379 (2007).

Acknowledgements

Working on a research project and writing a thesis requires technical, scientific and moral support from many people. I would like to thank and express my gratitude to everyone who helped me on this journey, directly and indirectly, and contributed to the great working environment in our research group:

- Dr. Jurgen Smet for continuously supporting me in many different ways. He provides guidance combined with scientific freedom which allows his group members to follow their own ideas. He ensures that his group has access to excellent equipment and working conditions. I learned a lot from him, also apart from science, having the chance to work on versatile tasks. He always trusted and motivated me.

- Prof. Dr. Peter Michler and Prof. Dr. Hans Peter Büchler for completing the PhD committee as co-examiners.

- All "Smetions" I had the pleasure to work with and learn from: Federico Paolucci, Matthias Kühne, Youngwook Kim, Daniela Tabrea, Johannes Geurs, Sven Fecher, Benedikt Frieß, Yijin Zhang, Joseph Falson, Daniel Kärcher, Pranoti Kshirsagar, Dong Zhao, Ding Zhang and Wafa Rouabeh. We laughed a lot and had fun not only in the office or lab but also outside work. I made many friends from all over the world and enjoyed our trips to the US, China, Korea, UK and Poland.

- In particular, I want to thank Steffen Wahl, Gunther Euchner, Sanela Göres as well as Yvonne and Benjamin Stuhlhofer for their great and invaluable technical support.

- All facilities of the Max Planck Institute which allow us researchers to focus on science and always provide a helping hand and guidance when needed. Special thanks to the whole nanostructuring lab team, and in particular to Marion Hagel and Achim Güth.

- Prof. Dr. Klaus von Klitzing for his continuous support of our group, inspiring discussions as well as for hosting great social gatherings in his garden or at castle Ringberg.

- Ruth Jenz for assisting us with the administrative work in her always helpful and kind way.

- All scientific collaborators, in particular the team around Robert Rölver and Theresa Lutz from Robert Bosch Central Research for our joint project on graphene based Hall sensors. Moreover, I want to thank Prof. Dr. David Goldhaber-Gordon and his group for having the chance to visit them at Stanford University and learn more about sample fabrication.

- My friends for distracting me from work and cheering me up when necessary.

- Anja - for your love, patience and support through very difficult and also dark times. I am looking forward to our common future and all small and big steps we are going to make together.

- My parents who always supported and believed in me from my childhood until this day. I cannot express the gratefulness, appreciation and love I have for my mother who I tragically lost before being able to finish this work. This is for you. *Non est ad astra mollis e terris via.*

List of publications

Y. Kim*, **P. Herlinger***, T. Taniguchi, K. Watanabe, and J. Smet, "Reliable Postprocessing Improvement of van der Waals Heterostructures," *ACS Nano* 13, 14182-14190 (2019). (* equal contribution)

D. Wolf, D.-H. Chae, T. Utikal, **P. Herlinger**, J. Smet, H. Giessen, and M. Lippitz, "Spectroscopy of Graphene at the Saddle Point," Book chapter in *Optical Properties of Graphene*, World Scientific, January 2017 (ISBN: 978-981-3148-74-1).

N. Chejanovsky, M. Rezai, F. Paolucci, Y. Kim, T. Rendler, W. Rouabeh, F. de Oliveira, **P. Herlinger**, A. Denisenko, S. Yang, I. Gerhardt, A. Finkler, J. Smet, and J. Wrachtrup, "Structural Attributes and Photodynamics of Visible Spectrum Quantum Emitters in Hexagonal Boron Nitride," *Nano Letters* 16, 7037 (2016).

Y. Kim, **P. Herlinger**, P. Moon, M. Koshino, T. Taniguchi, K. Watanabe, and J. Smet, "Charge Inversion and Topological Phase Transition at a Twist Angle Induced van Hove Singularity of Bilayer Graphene," *Nano Letters* 16, 5053 (2016).

G. Hassink, R. Wanke, I. Rastegar, W. Braun, C. Stephanos, **P. Herlinger**, J. Smet, and J. Mannhart, "Transparency of graphene for low-energy electrons measured in a vacuum-triode setup," *APL Materials* 3, 076106 (2015).

Schriftliche Bestätigung der eigenständig erbrachten Leistung gemäß § 6 Absatz 2 der Promotionsordnung der Universität Stuttgart

Die eingereichte Dissertation zum Thema „High Quality Graphene for Magnetic Sensing“ stellt meine eigenständig erbrachte Leistung dar.

Ich habe ausschließlich die angegebenen Quellen und Hilfsmittel benutzt. Wörtlich oder inhaltlich aus anderen Werken übernommene Angaben habe ich als solche kenntlich gemacht.

Die Richtigkeit der hier getätigten Angaben bestätige ich und versichere, nach bestem Wissen die Wahrheit erklärt zu haben.

Stuttgart, den 06.02.2022

Erklärung zur Übereinstimmung der digitalen Version mit der vorgelegten Printversion der Dissertation

Hiermit erkläre ich, Sven Patrick Herlinger, geboren am 30.06.1986, dass das von mir gereichte pdf-Dokument zur Dissertation mit dem Thema „High Quality Graphene for Magnetic Sensing“ in Inhalt und Wortlaut der ebenfalls eingereichten Printversion meiner Dissertationsschrift entspricht.

Stuttgart, den 06.02.2022