Universität Stuttgart

Quantum Randomness Certified By Different Quantum Phenomena

Von der Fakultät 8 Mathematik und Physik der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von

Xing Chen
aus Xi'an, China

Hauptberichter:	Prof. Dr. Jörg Wrachtrup
Mitberichter:	Prof. Dr. Stefanie Barz
Tag der mündlichen Prüfung:	07.06.2023

3. Physikalisches Institut der Universität Stuttgart
2023

Zusammenfassung

Die Erzeugung von Quanten-Zufallszahlen nutzt Quantenprozesse, die den Kollaps eines Superpositionszustands bei der Durchführung einer Messung beinhalten. Bei einem Quantenprozess ist das Messergebnis grundlegend unvorhersehbar, was zu echter Zufälligkeit in den erzeugten Zahlen führt. Wir bezeichnen diese Art von Zufallszahlengenerator als Quanten-Zufallszahlengenerator (QRNG), da Quantenprozesse an der Erzeugung von Zufallszahlen beteiligt sind.

Der häufigste QRNG ist der photonische QRNG. Bei dieser Art von QRNG gelangen Photonen aus einer Laserquelle in einen Strahlteiler. Nach dem Strahlteiler befinden sich die Photonen in einem Superpositionszustand aus reflektiertem Pfad und übertragenem Pfad. In jedem Pfad befindet sich ein Detektor, der als Messgerät fungiert. Bei der Durchführung einer Messung kollabiert ein Photon zufällig in einen Detektor, was zu einem Klick im Detektor führt. Der Klick im übertragenen Detektor wird als Rohbit 0 zugewiesen, und der Klick im reflektierten Detektor wird als Rohbit 1 zugewiesen. Idealerweise ist jede erzeugte Zufallszahl aus diesem QRNG eine Quanten-Zufallszahl. In der realen Welt ist jedoch die Zufälligkeit in den erzeugten Zufallszahlen keine reine Quantenzufälligkeit, da sie auch andere technische Ursachen als die Quantenmechanik haben kann. Zum Beispiel können die Klickereignisse auf den beiden Detektoren von Dunkelzählungen herrühren, die als klassisches Rauschen betrachtet werden.

Wir müssen einige Quantenphänomene nutzen, die klassisch nicht erklärt werden können, um die Quantennatur der Rohbits zu beweisen und sicherzustellen, dass die Zufallszahlen aus dem QRNG alle durch Quantenprozesse anstatt durch unerwartete klassische Störungen erzeugt werden. Nachdem die Quantennatur nachgewiesen ist, können Zufälligkeitszertifizierungsprotokolle basierend auf dieser Quantennatur formuliert werden, um die Entropie der Zufälligkeit zu quantifizieren.

Das Ziel dieser Arbeit ist es, unsere Fortschritte bei der Entwicklung von Zertifizierungsprotokollen für Zufälligkeit bei QRNGs durch Nutzung verschiedener Quantenphänomene darzustellen, um die Quantennatur der erzeugten Zufallszahlen sicherzustellen. Zu diesen Quantenphänomenen gehören der Einzelphotonen-Antibunching-Effekt, die Welle-Teilchen-Dualität in einem verzögerten Wahl-Experiment, die Nichtlokalität in einem Bell-Test und die nichtnull-dimensionale Zeugen von Quantenmessungen.

Im ersten Ansatz wird ein einzelphotonenbasiertes QRNG auf Basis eines Stickstoff-Fehlstellen-Zentrums implementiert und drei verschiedene Zertifizierungsprotokolle für Zufälligkeit entwickelt, um die Quantenzufälligkeit in den Rohdaten zu zertifizieren. Im ersten Modell werden alle experimentellen Ereignisse als Rohbits zur Extraktion von Zufälligkeit verwendet, und die Geschwindigkeit der Zufallsausgabe beträgt $5,10 \times 10^4$ Bit pro Sekunde. Im zweiten Modell werden nur Einzelphotonenereignisse als Rohbits betrachtet, und die Geschwindigkeit der Zufallsausgabe beträgt $4,74 \times 10^4$ Bit pro Sekunde. Im dritten Modell werden nur Tupelerkennungsereignisse unterhalb der Einheitslinie als Rohbits betrachtet, und die Geschwindigkeit der Zufallsgenerierung beträgt $34,37$ Bit pro Sekunde. Von diesen erreicht das zweite Protokoll, das den Einzelphotonenantibunching-Effekt nutzt, einen quellenunabhängigen Zufallszahlengenerator, ohne die Geschwindigkeit der Zufallsausgabe zu beeinträchtigen, was es zu einer idealen Wahl für einzelphotonenbasierte QRNGs macht.

Die zweite Methode konstruiert ein QRNG auf Basis eines verzögerten Wahl-Experiments ohne die Annahme einer fairen Stichprobe. Mit Hilfe der Wellen-Teilchen-Dualität stellt das Modell sicher, dass Photonen in überlagerten Zuständen an Detektoren ankommen und somit die Notwendigkeit einer fairen Stichprobe entfällt. Durch Anwendung dieses Modells auf ein verzögertes Wahl-Experiment [1] können wir 1.124 gleichmäßig verteilte Zufallsbits pro Sekunde erzeugen.

Der dritte Ansatz zertifiziert Quantenzufälligkeit aus schlupflochfreien Bell-Testdaten unter Verwendung von Bells Theorem [2] und dem Fernzustandsvorbereitung (RSP)-Dimensionenzeuge [3]. Das Modell des RSP-Dimensionenzeugen erhöht die Geschwindigkeit der Zufallsausgabe von $2,54$ Bit pro Tag auf $40,63$ Bit pro Tag und markiert einen wichtigen Schritt in Richtung praktischer Einsatz von Bell-Tests in der Zufallsgenerierung.

Schließlich wird ein QRNG basierend auf einem Kernspinsystem in einem NV-Zentrum untersucht, einschließlich zwei Zertifizierungsprotokollen für Zufälligkeit. Das erste Protokoll ist eine direkte Anwendung des W_2-Modells aus [4], und Zufälligkeit kann mit einer Geschwindigkeit von $0,87$ Bit pro Sekunde erzeugt werden. Im zweiten Dimensionszeugenmodell entwickeln wir ein Zertifizierungsprotokoll für Zufälligkeit basierend auf einem dreidimensionalen Dimensionszeugen W_3, und dessen Geschwindigkeit der Zufallsausgabe beträgt $1,33$ Bit pro Sekunde, das ist um 53% höher als $0,87$ Bit pro Sekunde.

Durch die Nutzung dieser Quantenphänomene tragen wir zur wachsenden Notwendigkeit sicherer, hochwertiger Zufallszahlen in verschiedenen Bereichen bei, einschließlich Kryptographie, wissenschaftlicher Simulationen und Algorithmusentwicklung.

Summary

Quantum random number generation utilizes quantum processes, which involve the collapse of a superposition state upon performing a measurement. In a quantum process, the measurement outcome is fundamentally unpredictable, resulting in true randomness in the generated numbers. We refer to this type of random number generator as a quantum random number generator (QRNG) since quantum processes are involved in generating random numbers.

The most common QRNG is the photonic QRNG. In this kind of QRNG, photons from a laser source go into a beamsplitter. After the beamsplitter, the photons are in a superposition state of reflected path and transmitted path. In each path, there is a detector acting as a measurement device. When a measurement is performed, one photon collapses into one detector randomly, resulting in a click in the detector. The click in the transmitted detector is assigned as raw bit 0, and then the click in the reflected detector is assigned as raw bit 1. Ideally, from this QRNG, each random number generated is a quantum random number. Still, in the real world, the randomness in the generated random numbers is not pure quantum randomness since it can have other technical causes other than quantum mechanics. For example, the click events on the two detectors can come from the dark counts, which are considered to be classical noise.

We need to utilize some quantum phenomena, which cannot be explained classically, to prove quantumness in the raw bits to guarantee that the random numbers from the QRNG are all generated by quantum processes instead of some unexpected classical noises. After the quantumness is proved, randomness certification protocols based on this quantumness can be formulated to quantify the entropy of the randomness.

This thesis aims to present our progress in constructing randomness certification protocols for QRNGs by leveraging different quantum phenomena to ensure the quantumness of generated random numbers. These quantum phenomena include the single-photon antibunching effect, the wave-particle duality of a delayed-choice experiment, non-locality in a Bell test, and nonzero dimension witness of quantum measurements.

In the first approach, a single-photon QRNG based on an nitrogen-vacancy (NV) center is implemented, and three different randomness certification protocols are built to certify quantum randomness in the raw data. In the first model, all the experimental events are used as raw bits to extract randomness, and the randomness output speed is 5.10×10^4 bits per second. In the second model, only single photon events are considered as raw bits, the randomness output speed is 4.74×10^4 bits per second. In the third model only tuple detection events below the unity line are considered raw bits, and the randomness generation speed is 34.37 bits per second. Among them, the second protocol, utilizing the single-photon antibunching effect, achieves a source-independent random number generator without compromising the randomness output speed, making it an ideal protocol for a single-photon QRNG.

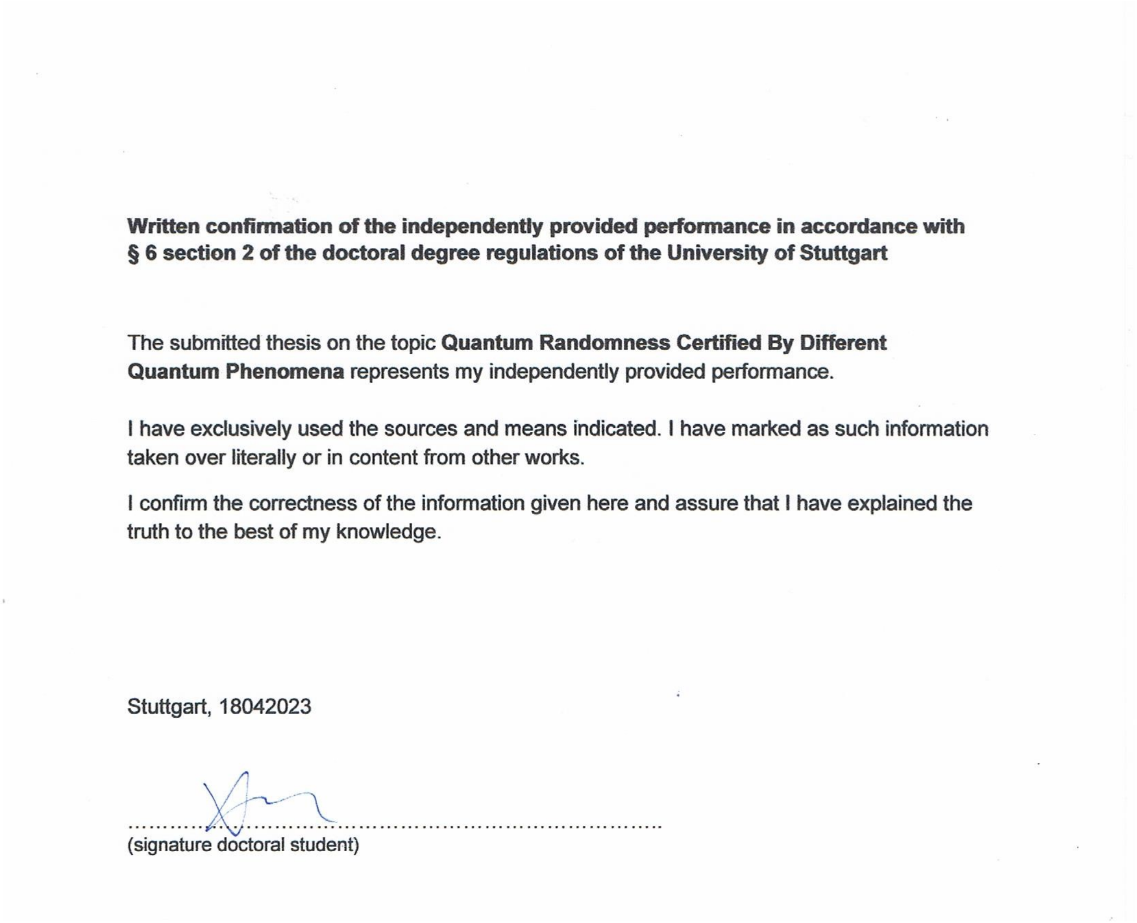
The second method constructs a QRNG based on a delayed-choice experiment without the fair sampling assumption. Using wave-particle duality, the model ensures photons arrive at detectors in superposition states, eliminating the need for fair sampling. By applying this model to a delayed-choice experiment [1], we can obtain $1,124$ uniformly distributed random bits per second.

The third approach certifies quantum randomness from loophole-free Bell test data using Bell's theorem [2] and remote state preparation (RSP)-dimension witness [3]. The RSP-dimension witness model significantly increases the randomness output speed from 2.54 bits per day to 40.63 bits per day, marking an important step towards the practical use of Bell tests in randomness generation.

Lastly, a QRNG based on a nuclear spin system inside an NV center is studied, including two randomness certification protocols. The first protocol is a direct application of the W_2 model from [4], and randomness can be generated with a speed 0.87 bits per second. In the second dimension witness model, we develop a randomness certification protocol based on a three-dimensional dimension witness W_3 and its randomness output speed is 1.33 bits per second, which is 53% higher than 0.87 bits per second.

By harnessing these four different quantum phenomena, we contribute to the growing need for secure, high-quality random numbers in different fields including cryptography, scientific simulations, and algorithm development.

Declaration



Written confirmation of the independently provided performance in accordance with § 6 section 2 of the doctoral degree regulations of the University of Stuttgart

The submitted thesis on the topic **Quantum Randomness Certified By Different Quantum Phenomena** represents my independently provided performance.

I have exclusively used the sources and means indicated. I have marked as such information taken over literally or in content from other works.

I confirm the correctness of the information given here and assure that I have explained the truth to the best of my knowledge.

Stuttgart, 18042023

……………………………………………………………………………

(signature doctoral student)

Contents

Abbreviations

BS	Beam-splitter
BSM	Bell state measurement
CHSH	Clauser-Horne-Shimony-Holt
DI	Device-independent
EOM	Electro-optic modulator
HBT interferometer	Hanbury Brown–Twiss interferometer
KS	Kochen-Specker
LHV	Local hidden variable
NIST	National Institute of Standards and Technology
NV	Nitrogen-vacancy
PRNG	Pseudorandom number generator
QRNG	Quantum random number generator
RSP-dimension witness	Remote state preparation dimension witness
SDI	Semi-device independent
WP	Wollaston prism

1 Introduction

Randomness is an important topic in both academic and industrial fields. On the one hand, research on randomness can deepen our understanding of nature [5, 6, 7]. On the other hand, randomness is an essential resource for cryptography [8, 9, 10, 11], algorithms, and scientific simulations. For example, the Monte Carlo method, a mathematical technique to numerically solve difficult and complex analytical problems using random numbers, finds applications in multiple fields, including physics, finance, biology, and chemistry. Randomness also plays a very crucial role in securing information. From the Caesar cipher to the Rivest-Shamir-Adleman (RSA) cryptosystem, the need for information encryption has a long history, and randomness is one indispensable element of encryption. With the advent of the information era, the need for information encryption is growing fast, and the demand for random numbers increases along with this growth. In this introduction, we discuss the history of randomness, the application of randomness, the motivation to go from pseudo-random numbers to quantum random numbers, and the development of QRNGs.

1.1 The history of randomness

Randomness has a surprisingly long history despite sounding like a very modern and scientific term. In ancient times, the concepts of randomness were often connected with fate or destiny. For example, around 7,000 to 8,000 years ago, religious shamans used marked objects such as fruit pits, seashells, and bones to tell people's future by trying to interpret the signs from the god. The so-called "signs" were distribution patterns by throwing those objects on a table or the ground. Each time the pattern was randomly formed and could not be repeated. The interpretation was also quite random, heavily dependent on shamans' personal tastes. Then later, around 5,000 years ago, dice were invented [12]. The invention of dice gave humans the ability to generate random numbers whenever needed–soon after its invention, randomness found its place in gaming [13] and even gambling [14].

Figure 1.1: Ancient roman die. This die is already very similar to the one used in daily life. Picture is adapted from [15]

However, what is randomness? The Greek philosophers, including Democritus, Aristotle, and Epicurus, discussed it in non-quantitative forms. For instance, in Aristotle's opinion, randomness is a genuine and widespread part of the real world and is subordinate to necessity and order [16].

Then for centuries, the argument about randomness stayed in philosophy aspect with no mathematical foundation. It was not until the 17th century that mathematicians developed another fundamental concept that could quantify randomness– probability– the most important mathematical property of random numbers. In 1654, Blaise Pascal corresponded with Pierre de Fermat, and in their exchange of letters, they established the major work for probability theory [17]. The work of Pascal and Fermat later influenced Leibniz's work on infinitesimal calculus, which in turn provided further development for the mathematical theory of probability and randomness. Later, the first textbook on probability theory was published in 1718 [18]. And then, the mathematical study of randomness continued to grow thereafter [19].

Probability only describes the mathematical nature of specific variables in a random sequence. When we want to quantify the randomness in the whole sequence, the term "entropy" is always followed. Entropy, a critical concept in studying randomness, was introduced by Rudolf Clausius in 1865 [20]. Later in 1877, Ludwig Boltzmann provided the mathematical definition of entropy $S = k_B \ln\Omega$, where k_B is the Boltzmann constant, and Ω is the number of different microstates with a given system energy. This definition interpreted entropy as a measure of the statistical disorder of a system. At the time entropy was defined, it had nothing to do with randomness. Randomness itself was still considered as a lack of knowledge of the system, since in Newton's classical mechanics, everything is deterministic with enough known input parameters. In other words, if all the forces acting on a system can be formulated with sufficient accuracy, it would be possible to make predictions of the state of such a system for an infinitely long time. This belief in the determinacy of nature was among almost all the scientists before the birth of quantum mechanics at the beginning of the 20th century, especially before the formulation of the Heisenberg uncertainty principle in 1927. But still, a mathematical description of randomness was missing.

It was not until 1948 that the connection between randomness and entropy was established– Claude Shannon's work in information theory gave rise to the entropy view of randomness for the first time [21]. In this perspective, randomness is considered as the opposite of determinism in a stochastic process. This is because if a stochastic system is deterministic, it has no randomness and is equivalent to saying its entropy is zero. Meanwhile, a nonzero entropy in the stochastic system means the system contains randomness. After this mathematical formulation of randomness, random numbers quickly found applications in various aspects of society, both in the industry and scientific fields [22].

During that time, tables of random numbers, usually generated using an electronic roulette wheel, were widely used in statistics and other scientific research [23]. Due to the growing need for random numbers, it was quite natural for researchers to turn to quantum mechanics for true randomness late 1950s [24, 25]. For example, in 1961, one of the first QRNGs based on radioactive process [25] was invented. Since then, the study of random numbers from the quantum process is mainly focused on the radioactive decay [26, 27]. In the 1980s, the first photonic implementation of QRNG was proposed by Morris [28], and later it was experimentally demonstrated. Since then, the studies on quantum random number generation began to grow, as shown in Fig. 1.2.

1.2 The application of randomness

One primary application of randomness is in the Monte Carlo simulation. Monte Carlo simulation finds application in multiple fields, including physics, finance, biology, and chemistry [29]. Another important application of randomness is related to cryptography, the art of encrypting and protecting information.

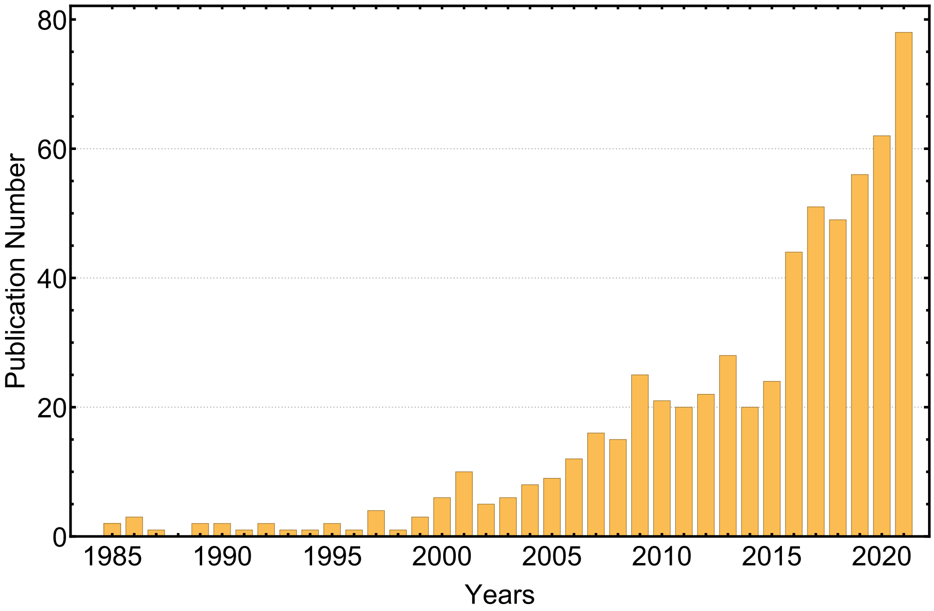


Figure 1.2: The publications about quantum random numbers. As we can see from this figure, the research on quantum randomness has been growing fast since the 1980s. Data is obtained from pubmed.ncbi.nlm.nih.gov

1.2.1 Randomness in scientific simulation

The importance of randomness in scientific simulation can be seen in Monte Carlo (MC)-simulation, which is a mathematical technique to numerically solve difficult and complex analytical problems by using random numbers [30]. In experimental particle physics, MC-simulation is important to design detectors since it can simulate the arrival of high-energy particles and help narrow the gap between theory and experiment. In astrophysics, especially in the simulation of galaxy evolution, MC-simulation plays a very important role [31].

In the mathematical field, the MC-simulation utilizes random numbers to solve various problems and observe the fraction of the numbers that obey some property or properties. This is very useful when the analytic solution to the problem is too complicated to be obtained. A straightforward example of using Monte Carlo methods is the approximation of the value of π, shown in Fig. 1.3.

Another powerful application for MC-simulation is numerical optimization. The problem is to minimize (or maximize) functions of some vectors that often have many dimensions. For instance, MC-simulation can be used to find a heuristic solution for the traveling salesman problem [32].

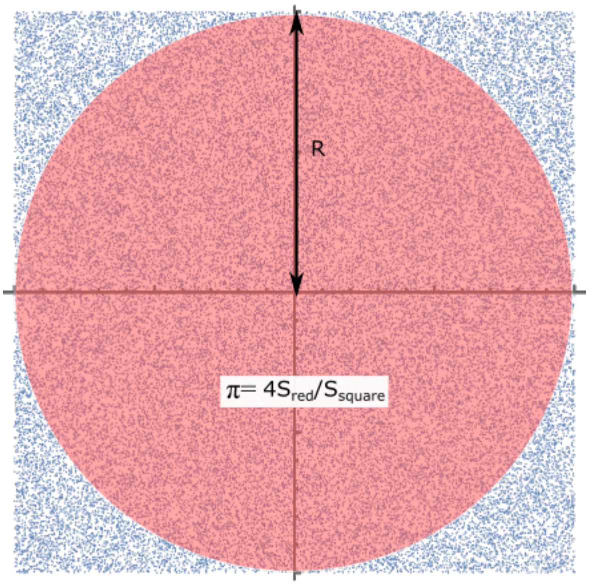


Figure 1.3: Monte Carlo simulation of π. The area of the red disc is $S_{\mathrm{red}} = \pi R^2$, and the area covered by the square is $S_{\mathrm{square}} = 4R^2$, so $\pi = 4S_{\mathrm{red}}/S_{\mathrm{square}}$.

Besides physics and mathematics, the MC-simulation method also has many applications in computer graphics, finance, business, and even law [33].

1.2.2 Randomness in cryptography

Since the invention of the electronic computer in 1945 [34], humans have begun to enter into the so-called information age [35]. From then on, information gains more and more significance in all aspects of human life: from information about one's house number to information regarding national security. Without exaggeration, in the 21st century, information has become one of the most important resources in the world. Big companies like Google, Amazon, and Facebook are eager to collect personal information to make more profits. In the meantime, governments are acting to protect private information in a more and more stringent way [36]. In most cases, randomness is indispensable to protecting information. The study of securing information techniques in the presence of eavesdroppers is called cryptography (from Ancient Greek means “hidden, secret").

The protection of information means not only secretly storing information but also communicating information privately. In fact, the need for secret communications existed long before the information age. For instance, the so-called Caesar cipher [37] was invented 2,000 years ago. In this cipher method, we take each letter of the message and replace it with the letter three positions after it in the alphabet, and the encrypted message will be unreadable to the enemy. The Caesar cipher only uses one random number (3 in the example) as the key to encrypt the information, and it is very easy to be hacked. Up to now, several modern cryptographic systems use much longer random numbers as the key to guaranteeing the encryption security of the message. For instance, in the Rivest-Shamir-Adleman (RSA) cryptosystem [38], the key sizes are as large as 1024 or 2048 bits [39]. In the Advanced Encryption Standard (AES) cryptosystem, we use 256 bits key size to implement the AES-256 to encrypt the message [40].

The keys in the abovementioned cryptosystems should be distributed between the message sender and the receiver so they can communicate smoothly. The distribution is usually not a problem for the RSA cryptosystem since it has a private decryption key and public encryption key. Its security is more guaranteed by the computation complexity rather than random numbers. While key distribution is the most challenging part of the AES method, the standard way is by using a sneakernet, courier service, or even a dead drop. Those ways are very slow, inefficient, and bear the danger of being intercepted by adversaries. Is there a way to do it more efficiently and securely? Quantum mechanics gives us the solution. More especially, quantum key distribution (QKD) offers us a non-hackable way to distribute keys [8]. There are several different QKD protocols [8, 41, 42, 43], but the basic idea behind them is the same: by using the laws of physics, such as the no-clone theorem, the distribution of the key between sender and receiver can be done in an unconditionally secure manner.

The crypto-systems we mentioned above, including RSA, AES, and QKD, use different ways to distribute the key. No matter how the user wants to distribute the key, the key needs to be created first. The key is a random sequence, which can be obtained from random number generators. In this process, the privacy of the keys must be guaranteed for cryptographic use. *Privacy* means that the keys are only known by legitimate users, such as the sender and targeted receivers. From random number generation to creating keys with privacy for cryptographic purposes is far from a trivial task.

1.3 Pseudo randomness versus true randomness

Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin— John von Neumann [44]

As mentioned above, since the advent of the information era, the applications of random numbers, especially in cryptography, have been growing quickly. This means we need to find ways to generate random numbers (such as keys for cryptography) fast and reliably. Since the birth of the first programmable and digital computer–ENIAC–computer-based random number generators have also emerged [45]. Such as the middle-square method proposed by J. von Neumann in 1946 [45]. Later, many algorithms-based random number generators were invented [46, 47].

Random numbers from algorithm-based random number generators are called pseudo-random numbers since their value is fully determined by an initial value (random seed). Correspondingly, those random number generators are called pseudo-random number generators (PRNGs).

The random numbers from PRNGs can be applied without problems in the MC-simulations, such as approximating the value of π. Such simulations do not require the security or the privacy of random numbers. But in the application of cryptography, there is a different story. There are many examples of insecure random numbers that lead to severe vulnerabilities. These include attacks on the SSL keys generated in the Netscape browser [48], attacks on the OpenSSL protocol [49], and the theft of Bitcoin due to the flaw of PRNG in Android phones [50]. The incentives for breaking a cryptographic system are very obvious: making money or fetching private information. Thus, using the best possible keys to secure our systems is an essential step in the protection of our information, not only personal but also public.

PRNGs are still in use in most user cases, mainly because they are easy to implement [46]. Nevertheless, we must pay special attention to some of the most relevant features of PRNGs when using them. The most important one is the seed problem. The seed fully determines the random number sequence from a PRNG: thus, if a PRNG is generated with the same seed, the same sequence of numbers will be produced; therefore, the security of the seed must be guaranteed. Another crucial thing about the PRNG is the finite length problem. The random numbers from any PRNG have a finite length, which means they will repeat themselves after a finite length. For some simulations, such as the simulation of the value of π, this repeat is not a problem as long as the pseudo-random strings are long enough for the simulations. However, this intrinsic period makes pseudo-random numbers vulnerable not only in the cryptography application, such as in the encryption of bank accounts, E-mail accounts, and all the encrypted coins [50] but also in some scientific simulations [51].

Even random numbers generated from coin tossing, drawing numbered balls, or using mechanical devices are, in principle, deterministic, as the outcomes of these physical processes can be predicted by classical mechanics. This includes chaos, as its results are deterministic since it is described by classical mechanics [52]. Consequently, random numbers from classical processes do not contain true randomness because all classical processes are considered deterministic [53].

Because of the fundamental limitations of random numbers from classical processes, QRNGs, whose randomness is guaranteed by quantum mechanics, are getting more and more attention.

Before we go into QRNG, there is one fundamental question: Does true randomness exist in the universe? Randomness can be understood as a lack of complete information of a number sequence. True randomness means that the random number cannot be predicted by anyone with any extent of knowledge of the random number generators. From this perspective, true randomness does not exist in the classical world, which is dominated by deterministic Newtonian dynamics. While in quantum mechanics, the situation is quite different. There are processes such that specific outcomes have a probability. For instance, in the Stern–Gerlach experiment [54], the electron spin will be deflected up or down randomly, and this cannot be predicted no matter how much information we know about the system. Of course, this could have other interpretations other than quantum mechanics, such as the many-worlds interpretation [55] and the Superdeterminism [56]. But all these interpretations are highly speculative. The interpretation from quantum mechanics is most preferred. Thus, in order to generate true randomness, quantum processes must be utilized.

The Born rule guarantees the unpredictability of random numbers from quantum mechanics. The Born rule describes the outcome of a quantum measurement as fundamentally probabilistic [57]. It states

> If the system is in a state $\Psi \in \mathbf{H}$, then the probability $P(a=\lambda_i \mid \Psi)$ that the eigenvalue λ_i of a is found when a is measured is
> $P(a=\lambda_i \mid \Psi) = |(e_i, \Psi)|^2$.
> In other words, if $\Psi = \sum_i c_i e_i$ (with $\sum_i |c_i|^2 = 1$), then $P(a=\lambda_i \mid \Psi) = |c_i|^2$.

H is the Hilbert space of the state Ψ, a is a quantum mechanical observable (i.e., position, momentum, polarization).

So the randomness from the measurement of a quantum superposition state is guaranteed by the correctness of quantum mechanics. As shown in Fig. 1.4, in a photonic QRNG, a photon shot into a beamsplitter can either be reflected or transmitted. The photon is in the superposition of these two paths until it has been detected (i.e., measured) by the two detectors, and the click events in the two detectors are truly random.

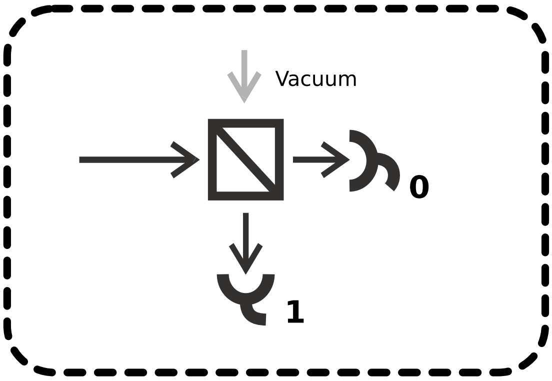


Figure 1.4: The scheme of photonic QRNG. A photon is shot into the beamsplitter. After the beamsplitter, it will be in the superposition state of the transmitted or reflected path until it hits the detectors.

1.4 Testing randomness

With so many applications of random numbers, one thing that needs to be figured out is how to quantify the quality of the random numbers. The very first thing we need to check is frequency. Based on the frequency stability [58], a sequence is considered random if two conditions are satisfied. (I) Use a function $f(n)$ to count the number of ones (or zeros) in a sequence of length $|n|$, and $\lim_{n\rightarrow\infty} f(n)/|n| = p$; (II) Use a function $g(m)$ to count the number of ones in a subgroup sequence $|m|$ of the original sequence, and $\lim_{m\rightarrow\infty} g(m)/|m| = p$. The second condition helps avoid considering sequences such as $10101010101010...$ as random sequences.

From the frequency perspective, the National Institute of Standards and Technology (NIST) developed the NIST Statistical Test Suite [59], which does not only the frequency test to a given sequence but also some other tests relevant to the frequency, including the linear complexity test, the serial test, and the random excursion test, etc.

NIST Statistical Test Suite alone cannot guarantee the privacy of random numbers. For example, for the number $\pi = 3.14159265358979323846264338327950288419716939937...$, which is a mathematical constant, and its decimal representation never ends. Apparently, π is a deterministic string, while from the perspective of the NIST test suite in Fig. 1.5, it is a perfect random sequence. This inadequacy makes it insecure to use random sequences, which can pass the NIST test, in cryptography [8, 60, 61, 62]. The so-called Kolmogorov complexity can address this loophole. Kolmogorov complexity of a sequence is the length of the shortest computer program (in a predetermined programming language) that can produce the sequence as output [63]. For a uniform random sequence r, its Kolmogorov complexity $K(r) \geq |r|$, where $|r|$ is the length of the random sequence. The sequence is not random if $K(r) < |r|$. Now, come back to π. π has an infinite length, while the computer program to reproduce it is quite short “a mathematical constant defined as the ratio of a circle’s circumference to its diameter". So from the perspective of Kolmogorov complexity, π is not a random sequence.

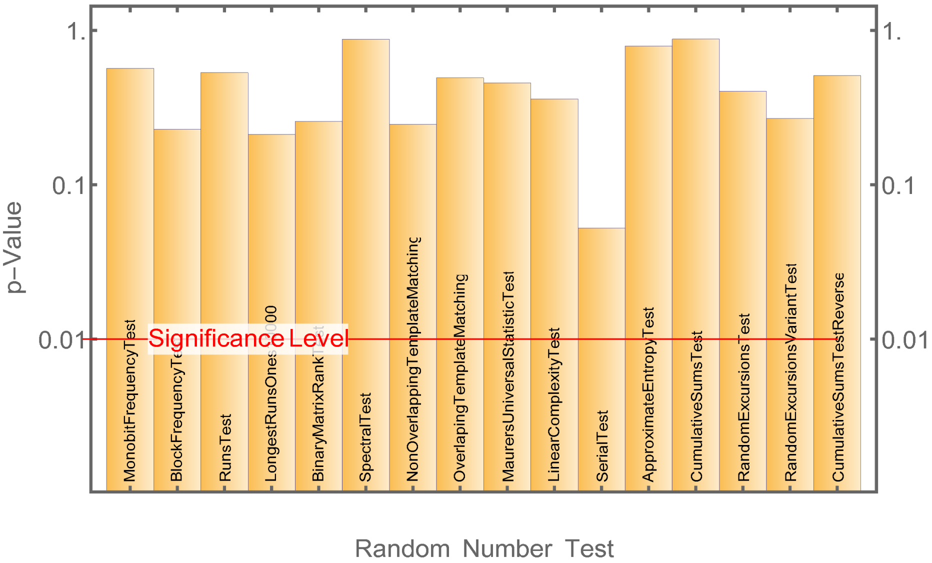


Figure 1.5: The NIST test suite for the first 10^5 digits in π. From the test results, we can conclude that the values of π have a uniform distribution.

For all the pseudo-random numbers from computer algorithms (or PRNGs), their Kolmogorov complexity is smaller than the final sequence. Thus, PRNGs cannot generate true randomness.

It seems that we can use this complexity to indicate the unpredictability of a sequence. However, for most sequences, their Kolmogorov complexity is unknown [63], and thus it cannot be used to prove the randomness of a sequence. This means that to describe a random sequence, it is insufficient to investigate the sequence alone, we must have knowledge about its source. In other words, we also need to know how the sequence is generated. For example, for random numbers from PRNGs, if we know how they are generated (i.e., the seeds are known), the sequence can be predicted deterministically, and the NIST test suite cannot guarantee their randomness. While for the random numbers from a quantum process, such as nuclear decay events, quantum mechanics guarantees that this process is unpredictable. In fact, quantum mechanical processes are believed to be the only known source of randomness in nature [64, 65, 66, 23], so the generation of random numbers by a quantum mechanical process in a QRNG is the desirable way to generate true randomness.

1.5 The development of QRNG

In the development of QRNGs, there are several kinds of QRNGs worth mentioning here. The first kind is the photonic QRNG. This kind of QRNG usually has the scheme in Fig. 1.4

In this scheme, photons are shot into a beamsplitter. After the beamsplitter, the photons are in a superposition state. The detectors in either path performing the measurements will break this superposition state and resulting in a random click. The photonic QRNG output speed relies on the single-photon detection technology, which is usually slow and expensive. For example, single-photon detectors from a provider like Thorlabs, etc., usually have a price of above 4,000 Euro, and the maximum count rate is often below 50MHz.

There is another category of photonic QRNG, which have a similar structure in Fig. 1.4 but can reach a much higher generation speed. This QRNG is based on continuous variables, for instance, vacuum fluctuations and phase noise [67, 68]. These fluctuations are detected by homodyne detectors, which can reach Gbps speed [69, 70, 71].

Because of the relatively mature technology of lasers and detectors, the photonic or similar QRNGs have already been commercialized for years [72]. However, for most of them, a user has problems verifying whether the random strings are truly from a quantum process. As shown in Fig.1.6, the photonic QRNG works normally in the left part of this figure, and we can get quantum random numbers from it. While in the right part of this figure, one of the detectors is misaligned and begins to receive white noise. In this case, if the count rate of this detector does not change, we cannot guarantee that the final output sequence contains randomness. One way to overcome this misalignment is to verify the source's legitimacy.

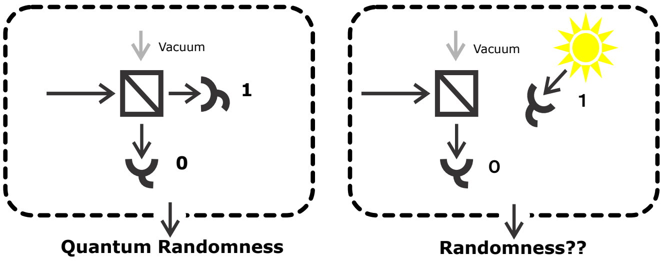


Figure 1.6: A well aligned and a misaligned QRNG. Due to the misalignment of the detector, the misaligned QRNG cannot generate quantum randomness.

Since this will reduce the dependency on the source, we call the source verifiable QRNG source-independent QRNG [65]. The basic idea of this type of QRNG is to use measurement results to characterize the source. In order to verify the source, the measurement setting needs to be switched randomly so that the source cannot predict which measurement comes next. In order to do so, a short random seed is required for switching the measurement settings. The advantage of this QRNG is obvious: Randomness can be generated even if the knowledge of the source is lacking or untrusted parties provide the source. In this scheme, the measurement devices are fully trusted and need to be well described. In contrast to the source-independent QRNG, there is also the so-called measurement-device-independent(MDI) QRNG. The source is well described in this kind of QRNG, while the measurement devices are untrusted. The advantages and disadvantages of this QNRG are just the reverses of the source-independent QRNG.

Both the source-independent and MDI-QRNG lift the dependency on the experimental device to some extent. However, they still require detailed characterization of the measurement device or the sources. When the experimental devices deviate from the theoretical models, the randomness can be compromised since realistic devices will inevitably introduce some unexpected classical noise that affects the purity and security of the quantum randomness. With the increasing security need for cryptography–which is highly based on random numbers–the security requirement for randomness generation is getting more and more demanding. In this background, the self-testing QRNG scheme is gaining more and more popularity. In the self-testing QRNG, the quantum randomness can be bounded independent of the experimental devices.

In order to realize the self-testing scheme, physical inequalities should be utilized [66]. One of the well-known inequalities is the Clauser–Horne–Shimony–Holt (CHSH) inequality [73]. The CHSH inequality was designed for a bipartite Bell test, shown in Fig. 1.7. The input x and y, output a and b from Alice and Bob form the CHSH inequality $S =$

$\sum_{a,b,x,y}(-1)^{a+b+xy}p(a,b \mid x,y) \leq 2$. In loophole-free Bell tests based QRNGs [66, 74, 75, 76, 77, 78], the random numbers can be generated without trusting any parts of the experimental devices. This kind of QRNG is called device-independent (DI) QRNG. The DI-QRNG has very demanding requirements on the experimental setup [61, 79, 80, 62], such as the locality and efficiency loopholes must be closed simutaneously [81]. The first loophole-free Bell test is realized in 2015 [61]. Because of the experimental challenge, the randomness output speed of the loophole-free Bell test based QRNG is extremely low, around a few bits per day [82, 62], which is far from practical use. So there is another kind of self-testing QRNG, which lift the stringent requirements on the experimental setup and have a relatively higher randomness generation speed. In this kind of self-testing QRNG scheme, the principle of quantum mechanics is trusted. Besides, some other general assumptions are made, such as the independence and memorylessness of the devices. Such a QRNG is also called semi-device-independent (SDI) QRNG since they rely on general assumptions about the experimental devices. The QRNGs based on dimension witness [83, 84, 4] and Kochen-Specker (KS) theorem [85, 86] are SDI QRNGs.

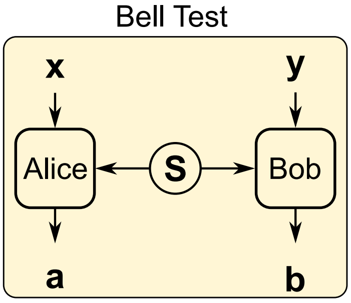


Figure 1.7: Scheme of bipartite Bell test. The two parties in a Bell test are traditionally named as Alice and Bob.

1.6 The structure of this thesis

The development of QRNGs will continue with the increased need for random numbers in the information era. This thesis makes its contribution to the development of QRNGs by performing experiments and formulating mathematical models for different quantum phenomena to generate quantum random numbers in a certifiable way. This thesis is structured as follows to give the readers a comprehensive understanding of our work.

In Chapter 2, four different quantum phenomena, including the single-photon antibunching effect, the wave-particle duality in a delayed-choice experiment, non-locality in a Bell test, and nonzero dimension witness of quantum measurement, are summarized. Each of these quantum phenomena shows quantumness in its unique way. We briefly discussed how to use them for randomness certification, and the main work in this thesis is that we build new randomness certification protocols for each quantum phenomenon mentioned in this chapter.

In Chapter 3, standard methods and tools for building a QRNG are introduced, including how to quantify the entropy of randomness with min-entropy, how to get a uniformly distributed random sequence with randomness extractors, and how the NIST Statistical Test Suite test the structure of a uniformly distributed random sequence. From Chapter 4 to Chapter 7, we performed experiment with each quantum phenomenon mentioned in Chapter 2 to generate quantum random numbers, and the corresponding QRNG protocols are illustrated in detail in each chapter.

In Chapter 4, the implementation of a single-photon QRNG is discussed, and the models to quantify the entropy in the raw bits are explained. As for single-photon sources, a large variety of single emitters as single-photon sources has been investigated in the past. One prominent example is the negatively charged NV center, a stable single-photon source [87]. Since few experiments were performed which utilize the single-photon emission of an NV center for quantum randomness generation, a detailed model to quantify the randomness from this single-photon QRNG is still missing. So we choose the NV based single-photon source to build a QRNG and build models to quantify the entropy of the random numbers from this QRNG.

This single-photon QRNG can be considered a source-independent random number generator, which does not require the trust of light sources. Moreover, our single-photon based source-independent QRNG does not require a random seed to change measurement settings to test the sources since the single-photon antibunching effect is utilized. However, it still needs to trust the measurement devices, especially the beamsplitter, since the randomness in this QRNG is guaranteed by the so-called fair-sampling assumption [73, 88, 89] on the beamsplitter to get quantum randomness.

In Chapter 5, we construct a QRNG model based on a delayed-choice experiment to get quantum random numbers without the fair sampling assumption. The fair sampling assumption assumes no post-selection of experimental events in the beamsplitter. The post-selection destroys the superposition of photon states before they reach the two detectors. We use the quantum phenomenon in a delayed-choice experiment to guarantee that photons arrive at the detectors in superposition states, so the fair sampling assumption is no longer needed. Then a mathematical model is built to quantify the entropy of randomness by utilizing the interference visibility in the delay-choice experiment.

In Chapter 6, we show how to certify quantum randomness from a loophole-free Bell test data by Bell's theorem [2] and RSP-dimension witness [83, 84, 4, 3]. With the Clauser-Horne-Shimony-Holt (CHSH) inequality [73] in Bell's theorem, the min-entropy of random numbers can be bounded in a DI-way [82, 66, 74, 75, 76, 77, 78]. Also, from the KS theorem [85, 86] and the dimension witness [83, 84, 4], SDI[1] QRNGs have been experimentally demonstrated. In all these DI and SDI QRNG approaches, the usage of fundamental physics inequalities makes it possible to quantify the entropy in the raw experimental bits without knowing the details of the experimental setup. In this chapter, we will apply both DI and SDI protocols in the same loophole-free Bell test data [62, 3] and build a Bell test QRNG in two different ways.

In Chapter 7, a QRNG based on a nuclear spin system is studied, including the corresponding QRNG model. The nuclear spin states inside an NV center are well isolated from the environment and can be operated at room temperature, and their state initialization and control are relatively easy [90, 91]. Such a stable quantum system has been used in the quantum computing area for years [92, 91, 93]. To our knowledge, the nuclear spin system has not been used to generate quantum randomness, so we built a QRNG with this quantum platform and quantified the entropy of randomness with two different dimension witness protocols.

Chapter 8 is the conclusion and outlook of this thesis. This chapter summarizes the main work in this thesis and points out the possible further development of each QRNG we presented in this thesis.

[1] For the SDI protocols mentioned in this thesis, the assumptions about the experimental devices are general, which means they are not supposed to characterize the devices in detail, and they do not belong to source-independent or measurement-device-independent protocols.

2 Randomness from different quantum phenomena

A basic scheme of QRNG is shown in Fig 2.1. From this figure, we know that, in order to build a QRNG, some basic steps are involved. First, we need to identify which quantum phenomenon can be used to guarantee the quantumness in the raw experimental data (raw bits). After identifying the quantum phenomenon, one performs an experiment to generate raw experimental data. Then, a mathematical model must be formulated to utilize the quantumness in the experimental data to quantify the entropy of randomness. With this mathematical model, the min-entropy [1] of randomness in the raw experimental data can be quantified. At last, with the value of min-entropy in the generated raw bits, randomness extractors must be used to extract uniformly distributed random sequences. The NIST Statistical Test Suite can test the quality of extracted random sequences.

In principle, quantum randomness is all from the same source–the unpredictability of the measurement results, which further comes from the collapse of a quantum superposition state. However, as mentioned in Fig 1.6, the randomness is not always guaranteed to be from the collapse of a quantum superposition state. In different quantum systems, the superposition state exists in different forms and can be utilized differently. Quantum randomness generated by such systems can be guaranteed or certified differently.

In this thesis, the term "certified" or "certification" refers to a process that utilizes quantum phenomena to produce random numbers and thus guarantees (certifies) that these numbers are genuinely random. In other words, the certified randomness comes from a purely quantum phenomenon without classical analogue [66]. The certification procedure involves specific protocols that can reveal certain quantum phenomena in the generated numbers, thus confirm the numbers were indeed generated using a quantum process. This chapter outlines four different quantum phenomena that can be utilized to achieve such a certification

[1] This concept, along with randomness extractors and NIST Statistical Test Suite, will be explained in detail in Chapter 3

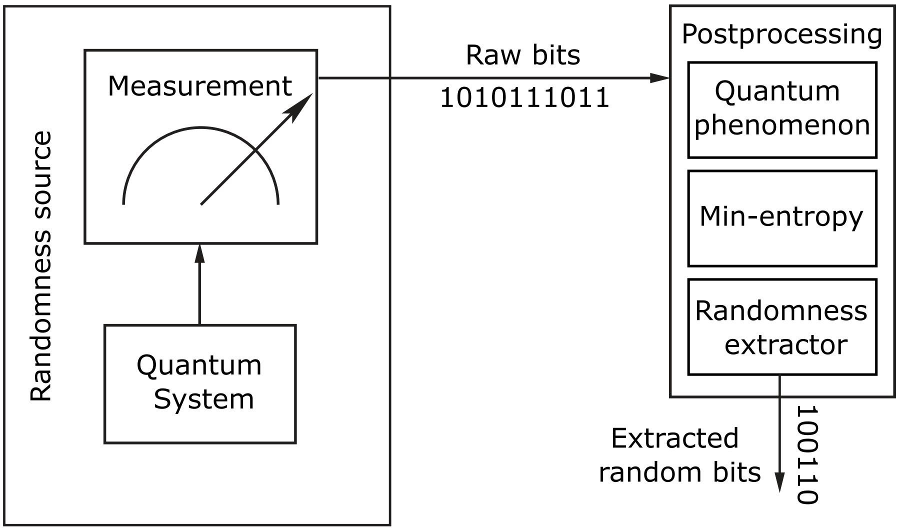


Figure 2.1: Basic scheme of a QRNG. A series of measurements are performed in a quantum system. The quantum system will randomly collapse into its eigenstates (usually two), and each eigenstate represents one raw bit (0 or 1). In the post-possessing stage, mathematical models are used to quantify the quantumness by the quantum phenomenon revealed in the raw bits, and this quantumenss can be used to bound the min-entropy in the raw bits. Then with the min-entropy value, a randomness extractor is implemented to extract uniformaly distributed random sequences.

procedure and enhance the authenticity and quantum nature of generating quantum random numbers.

2.1 Single-photon based QRNG

Due to the mature development of light sources such as laser and LED, most of the current QRNGs are based on the photonic system. These photonic QRNGs use different quantum superposition states to generate quantum randomness, including optical path [94], photon arriving time [95, 96], photon distributions [97, 98] etc. Here we focus on the optical path QRNG. This QRNG is very easy to understand, and the basic scheme is shown in Fig. 1.4.

This scheme is based on the behavior of single-photons at a beamsplitter. The beamsplitter can be represented in Fig. 2.2. Suppose the beamsplitter is 50/50, then the operation of the beamsplitter can be written as

$$|a\rangle \rightarrow \frac{1}{\sqrt{2}}(|c\rangle + i\,|d\rangle) \tag{2.1}$$

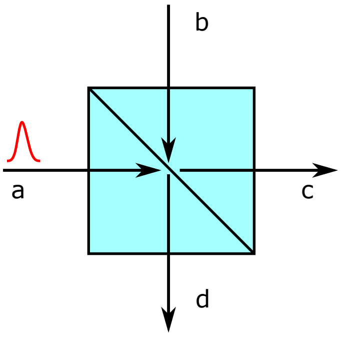


Figure 2.2: Action of a beamsplitter on a single-photon. The single-photon at input "a" will be in a superposition after it passes the beamsplitter.

The single-photon after the beamsplitter will be in a superposition state of $|c\rangle$ and $|d\rangle$. When performing measurement to this state with two detectors in path c and d, the two detectors both have a 50% chance to click. As shown in Fig 1.4, bit 1 is assigned to the reflected path d, and bit 0 is assigned to transmitted path c. If a continuous stream of photons is incident from port a, the click events in detectors c and d will register a random sequence like $0110001111100100100111000 10110....$ The min-entropy of this QRNG scheme depends mainly on the beamsplitter ratio.

Obviously, the beamsplitter is the central part of this QRNG because photons will only be in a superposition state after they pass the beamsplitter. Otherwise, the generated randomness could come from classical noise, as shown in Fig 1.6. With a regular laser, as we mentioned in the introduction, a user has difficulties verifying the legitimacy (including verifying the correct alignment and excluding classical noise in the click events of detectors) of the devices. The introducing of a single-photon source can overcome this problem.

2.1.1 Single-photon sources

Single-photons are nonclassical light, so classical light sources cannot simulate them. One of the major motivations for studying single-photon sources is the advent of quantum computing and quantum communication over the last few decades [99]. Since photons travel very fast (299,792,458 m/s in the vacuum) and interact weakly with their environment, single-photons are ideal photonic qubits for quantum computing. The information in quantum computing can be encoded in a quantum state of the photon using degrees of freedom, including but not limited to polarization and momentum energy. In quantum communication, most quantum key distribution protocols [8, 41, 42] have the best performance with single-photons, as more than one photon can comprise the security by allowing eavesdroppers to gain information about the key [100].

An ideal single-photon source can produce single-photons on-demand (a single-photon can be emitted at any time defined by the user), with a probability of emitting a single-photon being 100% and emitting multiple photons being 0%. However, achieving this ideal source is highly challenging in the real world due to inevitable losses and nonzero multiple-photon rates. There are two basic approaches to constructing single-photon sources [101]: One approach is by some isolated quantum systems that only emit one photon at a time. The other approach is by sources that emit photons in pairs, so the detection of one single-photon in one arm heralds the existence of a single-photon in another arm. Such isolated systems often include [102] color centers, quantum dots, single atoms, single ions, and single molecules. We often call them deterministic single-photon sources since single-photons can be produced with 100% probability in such isolated systems (in practice, the efficiency of photon collection is always less than 100% because of technique imperfection). In contrast, another approach, getting a single-photon by heralding, is called a probabilistic single-photon source since the generation of photon pairs is unpredictable. Single-photon sources such as parametric downconversion (PDC) in bulk crystals, waveguides, and four-wave mixing (FWM) in optical fibers belong to probabilistic single-photon sources.

Although the isolated quantum systems are from different materials in deterministic single-photon sources, their basic principle is very similar. When the user wants to generate a single-photon, an external laser is used to excite the system into a higher energy state, which will emit a single-photon when decaying to a lower energy state. This basic structure is shown in Fig. 2.3.

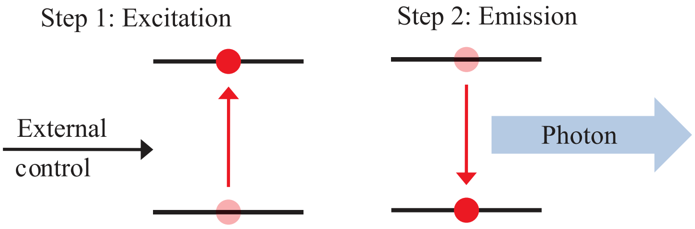


Figure 2.3: Single-photon emitting system for deterministic sources. This emitting system usually has two energy levels: one excited level and one ground level. The figure is adapted from [102].

For a probabilistic single-photon source, the basic structure is shown in Fig. 2.4.

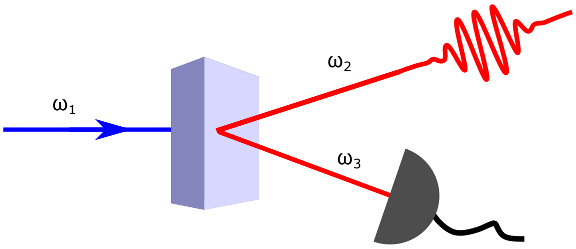


Figure 2.4: One common probabilistic single-photon source. This is a heralded single-photon source. By pumping a photon with frequency ω_1 into a nonlinear crystal, one photon pair with frequency ω_2 and ω_3 ($\omega_1 = \omega_2 + \omega_3$) will come out. The detection of one photon with frequency ω_3 will herald the existence of a single-photon with frequency ω_2.

Both approaches have their advantages and drawbacks from the perspective of current technology. Probabilistic sources often suffer from lower generation probability and multiple-photon pair generation. However, they are relatively easier to prepare and have been utilized in many applications [101]. While a deterministic source is often experimentally challenging, and they often suffer from lower indistinguishability [101]. However, with the development of technologies, the distinguishability and stability of deterministic single-photon sources continue to increase, and they are leading closer to ideal [103, 104]. In this thesis, we will talk more about one kind of deterministic source, which is based on the NV color center [87]. Before going into this, we first discuss how to qualify a single-photon source.

2.1.2 Antibunching of single-photons

Single-photon fidelity can be referred to as the absolute fidelity between the generated single-photon and an ideal single-photon state $|1\rangle_k$, where k defines the field mode of the single-photon (including spatial mode, continuous-wave mode, and pulsed temporal mode). In order to get this fidelity, quantum state tomography is needed. However, this procedure is difficult, time-consuming, and often not used in practice.

Note that an ideal single-photon source would emit a single-photon each time with zero probability of multiple-photon emission. The probability of detecting multiple-photon can be characterized by the second-order coherence function $g^{(2)}(\tau)$ [105, 106, 107], where τ is the time delay of one photon followed by another. For an ideal single-photon source, $g^{(2)}(0) = 0$, and $g^{(2)}(\tau) > g^{(2)}(0)$, since after the emission of one photon, the emitter (as shown in Fig: 2.3) must be excited again before a second photon can be emitted, and this takes time (usually around few ns). Theoretically speaking, the two-level single-photon emitter can never emit two photons simultaneously. In practice, dark counts of detectors, meta-stable energy levels within the two-level system, and external light contribute to $g^{(2)}(0)$, making it larger than zero. So $g^{(2)}(0)$ can be used as a criterion to judge the quality of a single-photon source. For comparison and completeness, from a laser source, the photons are emitted independently, and we will get $g^{(2)}(\tau)$=1. While for a thermal light source, one often has $g^{(2)}(\tau) < g^{(2)}(0)$. The $g^{(2)}(\tau)$ function is usually measured by the Hanbury Brown-Twiss (HBT) interferometer [108], which is shown in Fig. 2.5

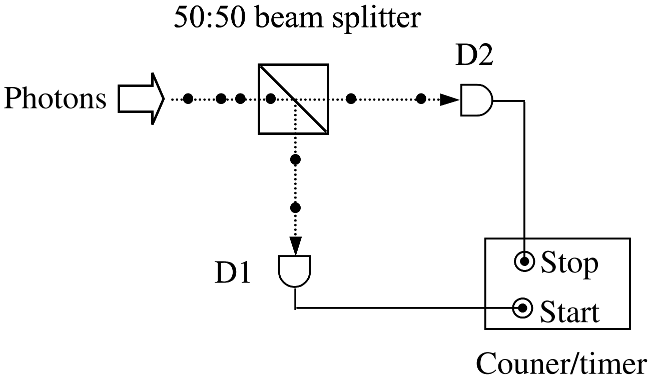


Figure 2.5: Hanbury Brown–Twiss interferometer. The pulses from the single-photon counting detectors D1 and D2 are fed into the start and stop inputs of an electronic counter/timer. The counter/timer counts the number of pulses from each detector and records the time that elapses between the pulses at the start and stop inputs. The figure is adapted from [109]

According to the value of $g^{(2)}(0)$, photon distributions can be categorized into three different kinds [109]:

- Bunched light $g^{(2)}(0) > 1$
- Coherent light $g^{(2)}(0) = 1$
- Antibunched light $g^{(2)}(0) < 1$

The illustration of these three different kinds of light is shown in Fig. 2.6.

Both bunched and coherent light have classical equivalents, which can be obtained in a classical way [109]. Antibunched light has no classical counterpart, representing a purely quantum phenomenon. This means that the light from a legitimate single-photon source cannot be simulated classically.

The antibunching curve of single-photons from our experiment is shown in Fig. 2.7, In order to get the antibunching curve, the setup shown in Fig. 2.5 is utilized. If the experimental

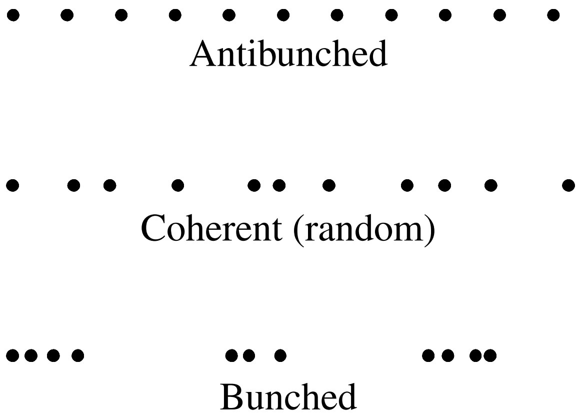


Figure 2.6: Comparison of different photon distributions. The distributions are from the antibunched, coherent, and bunched light. The figure is adapted from [109]

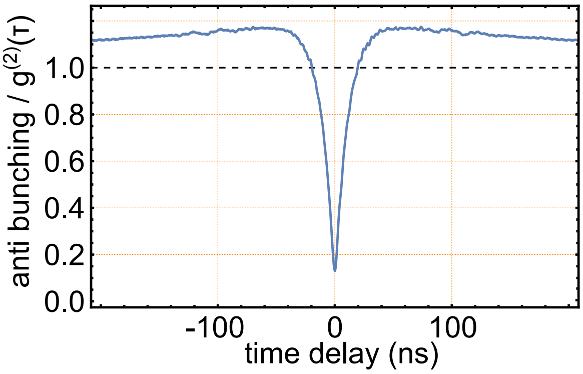


Figure 2.7: The single-photon antibunching curve. A small fraction of this antibunching curve is larger than 1. This is due to the extra meta-stable energy level inside the ideal two-level system, which will be discussed shortly. The experimental data is from our home-built NV single-photon source.

data reveals the antibunching effect (with $g^{(2)}(0) < 1$), this indicates that the clicks of the two detectors contain a quantum phenomenon that cannot be explained classically.

2.1.3 NV single-photon source

The NV center has been experimentally singled out since 1997 [110]. Besides its properties as a nanoscopic sensor and tool for spin-based quantum information processing, it represents a stable single-photon source with up to a few million counts per second [87, 111, 112]. A variety of single-photon based implementations was realized in the defect centers of diamond, for example, quantum cryptography [113], quantum computing [91] and other fundamental experiments [1].

The structure of the NV defect is shown in Fig. 2.8. The NV center can exist in negative

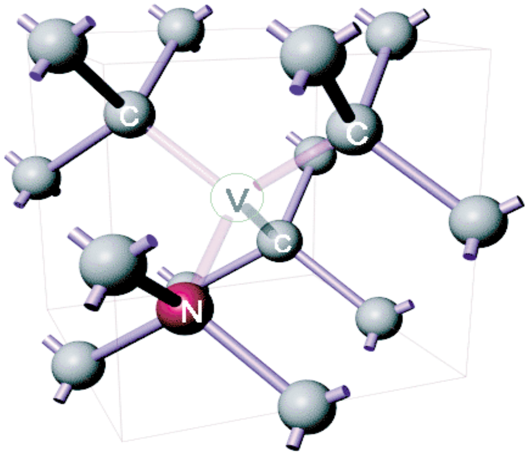


Figure 2.8: The structure of NV color center. The NV center is a point defect in the diamond lattice. It consists of one nitrogen atom adjacent to one atomic vacancy. The figure is adapted from [114]

charge state NV$^-$ and neutrally charged state NV^0. NV$^-$ and NV^0 have different zero phonon lines [115]. We use the NV$^-$ state to generate single-photons.

As discussed above, a deterministic single-photon source should be a two-level quantum system with one ground level and one exciting level. Since the ground level's excitation and the excited level's subsequent decay take a finite time, only one photon can be emitted at a time [87]. The NV$^-$ state can be described as a two-level system plus a metastable energy level in Fig. 2.9.

When the ground state $|g\rangle$ is pumped into the excited state $|e\rangle$, and $|e\rangle$ is decaying into $|g\rangle$, a single-photon will be emitted. In the energy system of NV$^-$, the excited state $|e\rangle$ is also thermally coupled with a metastable state $|s\rangle$. The metastable $|s\rangle$ is referred as a "shelving" state since the $|e\rangle$ to $|g\rangle$ emission ceases in this state. The existence of the shelving state $|s\rangle$ decreases the single-photon emission rate and causes photon bunching in the emitted single-photons. The photon bunching part can be observed in Fig. 2.7. This bunching part does not affect the non-classicality of the single-photon source since $g^{(2)}(0)$ is still less than 1 in this case.

With this NV based single-photon source, we build a QRNG. This QRNG uses the anti-bunching effect of single-photons to guarantee the source's validity and the photon path's well alignment. Then, quantum random numbers will be generated with the fair-sampling assumption on the beamsplitter and the detectors. The details are discussed in Chapter 4.

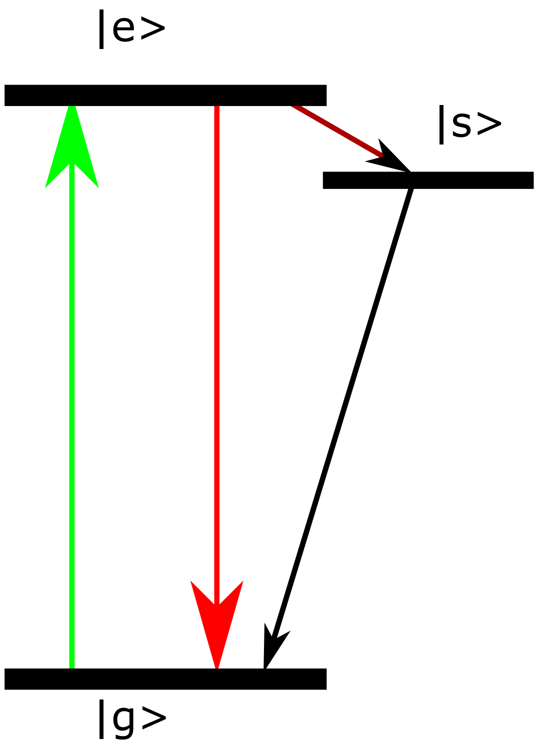


Figure 2.9: Three-level system of NV$^-$ state. The ground state is $|g\rangle$, the excited state is $|e\rangle$, and $|s\rangle$ is a metastable state.

2.2 Delayed-choice experiment based QRNG

The QRNG built on single-photon sources solves the problem of optical path misalignment and the legitimacy of the source. Nevertheless, the randomness generation is still based on the collapse of the superposition state after the beamsplitter. The fair-sampling assumption is applied to the beamsplitter and the two detectors to guarantee that the random numbers from this QRNG are from the collapse of superposition states. The fair-sampling assumption means that the experimental data is not post-selected to repeat desired results [73, 88, 89]. In other words, we trust that the beamsplitter and single-photon detectors function correctly.

In a single-photon based QRNG, the fair-sampling assumption is critical. It cannot be avoided since the antibunching of single-photons cannot tell whether the antibunching curve from the measurement results is post-selected. One basic fair-sampling strategy is explained in Fig. 2.10. When we do not assume the fair-sampling for the beamsplitter, it is equivalent to putting samplers after the two output paths of the beamsplitter. The samplers can block the photons coming out from the beamsplitter and use some pre-programmed strings to simulate the results. For example, suppose the detected events that come out of the beamsplitter are

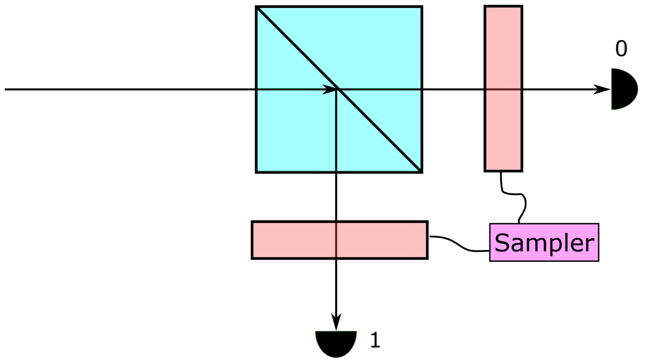


Figure 2.10: Sampler behind a normal beamsplitter. The samplers can block the photons coming out from the beamsplitter and use some pre-programmed strings to simulate the results.

$0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0...$, the sampler has a pre-programmed series $0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0$. This series can post-select the original series to match its sequence order at any time (in other words, the original series can only reach the user's detector when it matches with the pre-programmed series). For example, the original antibunching curve is shown in Fig. 2.7. In Fig. 2.11, we use one pseudo-random number sequence (for example, π) to post-select the measurement results, and the new curve looks very similar to the original one.

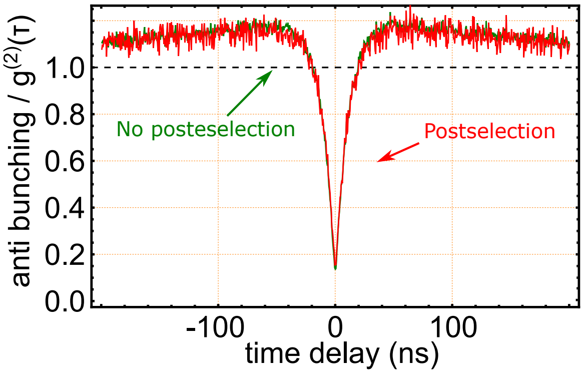


Figure 2.11: The post-selected antibunching curve. The post-selected experimental data still show the antibunching effect. This figure shows that the antibunching curve shows no statistical difference between no post-selection and post-selection cases.

Note that the post-selection destroys the superposition states of photons before they can reach the two detectors. In order to guarantee the quantumness in the generated random number, the superposition of photon states before reaching the detectors must be confirmed. Next, we illustrate how can a delayed-choice experiment give us the ability to check the status of photon superposition states after the beamsplitter and thus lift the fair-sampling assumption.

2.2.1 Double-slit experiment

We first introduce the double-slit experiment, which is highly relevant to the delayed-choice experiment.

In the 17th century, there were two theories about light: one explained light as waves, with the representative scientist being Huygens [116]; the other, from Newton, described light as a stream of fast particles. The latter received a more general acceptance because of Newton's authority in natural science. However, Young's double-slit experiment in 1801 [117] clearly demonstrated the wave behavior of light–wave interference patterns–as shown in Fig. 2.12. Later, with the performing of more similar experiments [118], the wave theory of light began to gain increasing support. However, in the year of 1905, Einstein put a stop to this trend. In his work [119], he successfully explained the photoelectric effect by assuming lights consist of "energy quanta which move without splitting and can only be absorbed or produced as a whole". By this time, scientists got confused about the behavior of light, and it seemed light could be both particle and wave. Nevertheless, no one brought it up until 1924, Louis de Broglie postulated that all particles, regardless of the mass, can behave as waves [120]. This postulation was later proven by experiments with electron and helium atoms [121, 122]. These experiments showed quantum-scale objects such as photons and electrons have wave-particle duality, and they cannot be simply described by classical concepts "particle" or "wave".

The wave-like behavior of a particle arises because, in quantum mechanics, all the information about a particle can be described by its wavefunction, and this function evolves according to Schrödinger equation. Furthermore, the particle-like behavior of the particle is related to measurement in quantum mechanics. The measurement performed on it will collapse its wave function, resulting in a peaked function at some location.

Based on the wave-particle duality of particles, the double-slit experiment was successfully performed with single-photons, electrons, atoms, and molecules [124, 125, 126]. In all the double-slit experiments, if the which-path information of every single particle is known, the interference pattern will not show. If the which-path information is not revealed, the interference pattern will show. This behavior can be understood from the perspective of quantum superposition. Take photons as an example: When photons travel through the two slits, each photon will be split into two different paths, and then each photon will be in the superposition state of the two paths. If which-path information is known, it means the photon

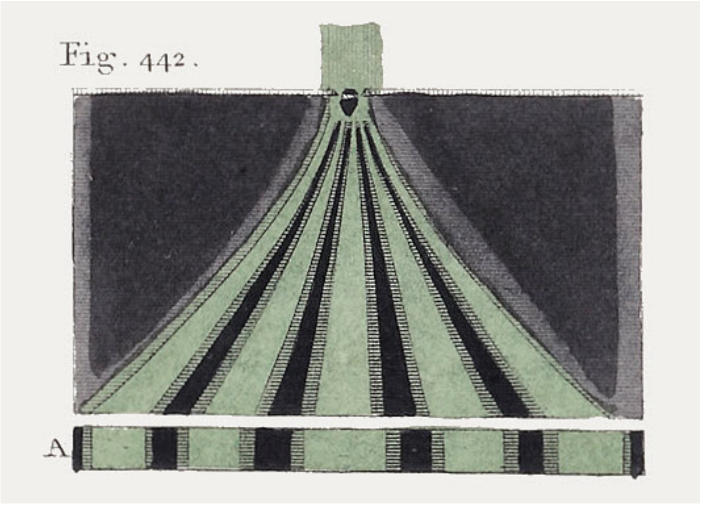


Figure 2.12: The interference pattern from Young's double-slit experiment. In a lecture by Young in 1802 to London's Royal Institution, the interference pattern of his double-slit experiment was shown. Figure is adapted from [123]

is measured before they reach the screen, and no interference pattern is shown on the screen. In this case, it is like photons traveling through only one slit at a time, showing the photons' particle-like character. If the photons are combined again after the two slits, the amplitude of the two superposition paths will interfere with each other, and the density distribution of the photons with different phase shifts will be shown on the screen, which is the wave-like character of the photons.

To give a more quantitative picture of this, we consider an equivalent experimental setup of the double-slit experiment–Mach-Zehnder interferometer (MZI), which was first proposed by Zehnder in 1891 [127] and was refined by Mach in 1892 [128].

2.2.2 Mach-Zehnder interferometer

The Mach-Zehnder interferometer is a device that is used to determine the relative phase shift of two paths. Both detectors can observe interference patterns with continuous phase shifts in one of the arms. In Fig. 2.13, a photon enters into the $\mathrm{BS}_{\mathrm{input}}$, and it will be in the superposition of path 1 and path 2. The superposition state of path 1 and path 2 will interfere at $\mathrm{BS}_{\mathrm{output}}$, and the detection probability in each detector will depend on the phase shift φ. Next, we derive the relationship between the detection probability of each detector and the phase shift φ.

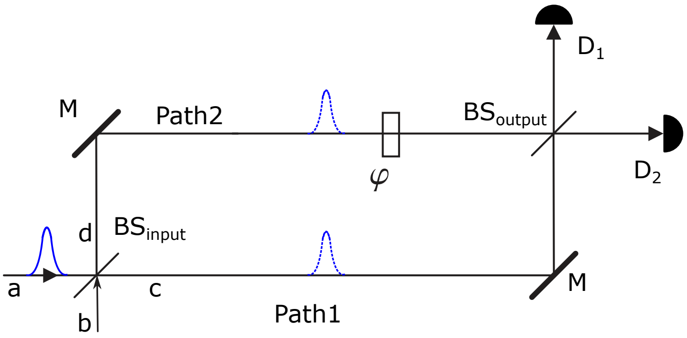


Figure 2.13: A schematic Mach-Zehnder interferometer.

In Fig. 2.13, port (b) is the vacuum state. Suppose the $\mathrm{BS}_{\mathrm{input}}$ is a 50/50 beamsplitter, and then photon incident via port (a) and have equal probability of being in port (c) or (d). This can be written as [129]

$$|a\rangle \xrightarrow{\mathrm{BS}_{\mathrm{input}}} \frac{1}{\sqrt{2}}(|c\rangle + i\,|d\rangle) \tag{2.2}$$

Then photon beams in port (c) and (d) travel long path 1 and path 2, and they will be superposed again after the $\mathrm{BS}_{\mathrm{output}}$, their quantum state evolves in the following way

$$\begin{aligned} |c\rangle &\xrightarrow{\mathrm{BS}_{\mathrm{output}}} \frac{1}{\sqrt{2}}(|e\rangle + i\,|f\rangle) \\ |d\rangle &\xrightarrow{\varphi} e^{i\varphi}\,|d\rangle \xrightarrow{\mathrm{BS}_{\mathrm{output}}} \frac{e^{i\varphi}}{\sqrt{2}}(i\,|e\rangle + |f\rangle) \end{aligned} \tag{2.3}$$

Considering $|c\rangle$ and $|d\rangle$ are in superposition state after $\mathrm{BS}_{\mathrm{input}}$, then we have

$$|a\rangle \xrightarrow{\mathrm{BS}_{\mathrm{input}}} \frac{1}{\sqrt{2}}(|c\rangle + i\,|d\rangle) \xrightarrow{\varphi} \frac{1}{\sqrt{2}}(|c\rangle + ie^{i\varphi}\,|d\rangle) \xrightarrow{\mathrm{BS}_{\mathrm{output}}} ie^{i\varphi/2}\left[-\sin\frac{\varphi}{2}|e\rangle + \cos\frac{\varphi}{2}|f\rangle\right] \tag{2.4}$$

From this, we can easily calculate the probability of finding photons in ports (e) and (f)

$$\begin{aligned} p_e &= \sin^2\frac{\varphi}{2} = \frac{1}{2}(1-\cos\varphi) \\ p_f &= \cos^2\frac{\varphi}{2} = \frac{1}{2}(1+\cos\varphi) \end{aligned} \tag{2.5}$$

With the change of phase-shift φ, the detection probability will vary in ports (e) and (f), and correspondingly this will change the count rate of each detector. In Fig: 2.14, we can see that the interference pattern is very similar to the interference pattern in a Young-type double-slit experiment.

The above derivation of detection probabilities with different phase-shift φ is based on normal beamsplitters, but the results are similar for polarizing beamsplitters, see appendix for the details.

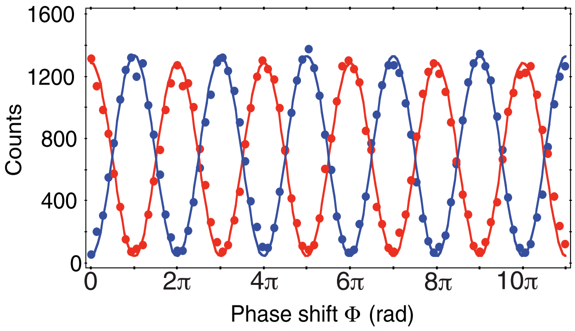


Figure 2.14: Interference pattern from a MZI setup. The interference pattern is similar to the one in a Young-type double-slit experiment. The figure is adapted from [1]

2.2.3 Wheeler's delayed-choice experiment

One important variant of the MZI experiment is Wheeler's delayed-choice experiment, which was one famous thought experiment in quantum physics. The original experiment was proposed in 1978 [130]. The scheme of this experiment is shown in Fig. 2.15. This experiment contains a MZI and a single-photon wave packet as input. This delayed-choice experiment was brought up to highlight the inherently nonclassical principle behind wave-particle duality.

In the top figure of Fig. 2.15, one single-photon wave packet is shot into a half-silvered mirror (labeled as $\frac{1}{2}S$ in the figure, the function is similar to a beamsplitter) from the left side. After the $\frac{1}{2}S$, there are two different possible paths $2a$ and $2b$. The single-photon can either go exclusively along one path (particle-like nature) or exist simultaneously in both paths (wave-like nature), depending on the status of the second $\frac{1}{2}S$. If the second $\frac{1}{2}S$ is removed from the setup, with perfect mirrors and detectors, one photon will trigger one click in one of the two detectors. After a while, one finds both detectors fire with equal probability, and their fire order is completely random. In this case, each photon has traveled only in one route, as pointed out by Wheeler, "[. . .] one counter goes off, or the other. Thus the photon has traveled only *one* route" [131], and the photon shows its particle nature. In the other case, if the second $\frac{1}{2}S$ is placed in the position as shown in the top figure, with identical path length (i.e., no phase-shift), only one detector (the bottom right one) will fire, the other one will never fire. In this case, the photon wave packet traveled in both routes and interfered in

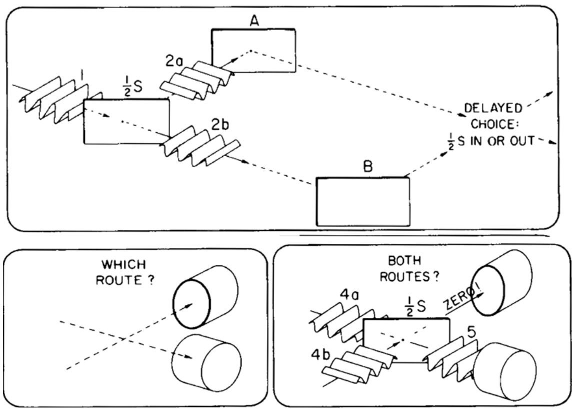


Figure 2.15: Wheeler's delayed-choice gedanken experiment with a single-photon in a Mach-Zehnder interferometer. Top: The half-silvered mirror($\frac{1}{2}S$) of the interferometer can be placed or removed. Bottom left: When the half-silvered mirror is removed, the photon path will be revealed by the click in the detectors. Bottom right: When the half-silvered mirror is placed in the interferometer, the detection probability of each detector depends on the length difference between the two arms. The mirrors A and B have 100% reflectivity, and the detectors have 100% detection efficiency. The figure is adapted from [131]

the second $\frac{1}{2}S$, causing constructive interference in one detector and destructive interference in another. This demonstrates the wave-like nature of the photon. In Wheeler's words, this is "...evidence that each arriving light quantum has arrived from both routes" [131].

From the perspective of quantum mechanics, wave-like nature means the photon is in the superposition state of both paths, and particle-like nature indicates that the photon is being detected before it can interfere with itself. In other words, the second $\frac{1}{2}S$ only decides whether the superposition state of photon paths interferes or not and does not determine which path the photon goes. With or without it, the photon always travels in two paths to reach two detectors. In order to prove the validity of the quantum mechanic explanation, Wheeler proposed a "delayed-choice" version of the experiment in Fig. 2.15, where the choice of removing the second $\frac{1}{2}S$ is made after the photon has passed the first $\frac{1}{2}S$. To illustrate how this will cause a counter-intuitive phenomenon, Wheeler proposed a most dramatic version of it: "delayed-choice gedanken experiment at the cosmological scale" [131]. The basic scheme is shown in Fig. 2.16

Considering the distance between the quasar and the receptor on Earth, the choice of removing or placing the second beamsplitter is made long after the photon enters into the cosmic interferometer (i.e., emission by the quasar). The exciting part of this delayed-choice experiment is that it rules out any causal influence from the emission of a photon to the decision, which might instruct the photon to behave as a particle or as a wave. In other words, the

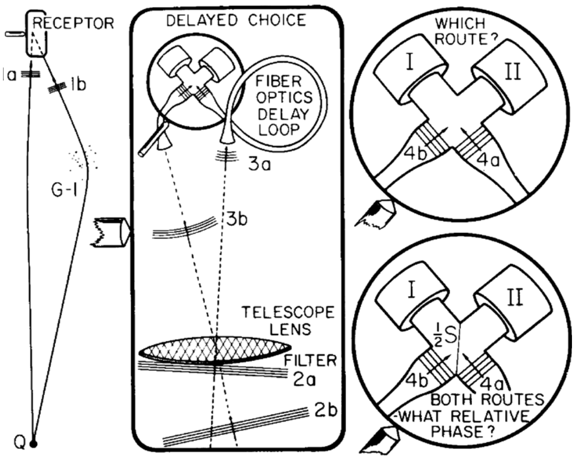


Figure 2.16: Wheeler's delayed-choice gedanken experiment at cosmological scale. In this experiment, Wheeler discussed what would happen if photons emitted by a quasar, which is millions of light-years away from Earth, pass a gravitational lens ($G - 1$). In a classical picture, when photons pass by the gravitational lens, each photon would have to "decide" whether to go one way around the lensing galaxy (behave as a particle) or go both ways around the lensing galaxy (behave as a wave). Then the photon arrives at the telescopes on Earth. Because of the gravitational lensing, the observers will see two pictures of the same quasar from two different paths. The two paths are observed separately in the upper right figure, and the photons behave as particles when arriving at the gravitational lens. Then if we direct the output of the two telescopes into a beamsplitter, one output will be very bright (indicating constructive interference), and the other output will be essentially zero, indicating that the incoming photons are behaving as waves. This means that photons retroactively decided to travel as waves when approaching the lens millions of years ago. The figure is adapted from [131]

spacelike separation in the cosmological scale excludes unknown communication from this decision to the choice of the receiver on Earth.

Because of the complexity of such an experiment at the cosmological level, the first loophole-free version of Wheeler's delayed choice experiment was realized in a "normal" scale by Jacques *et al.* [1]. In this experiment, they used the NV color center as the single-photon source, and the entry of the polarization interferometer and a fast (electrical optical modulator) EOM were separated by a 48-m-long fiber path. Their experiment was performed in 2007, and the experiment scheme is shown in Fig. 2.17.

The importance of this experiment is that it realized Wheeler's delayed choice experiment with a single particle quantum state (single-photon) and relativistic spacelike separation between the choice of interferometer configuration and the entry of the photon into the first beamsplitter of the interferometer. The space-time diagram of this experiment is shown in Fig. 2.18. The sequence for the measurement applied to the nth photon contains three steps. The first step is made by the QRNG (shown in blue), which creates a binary random number (0 or 1) to determine the configuration of the interferometer. Simultaneously, when the random

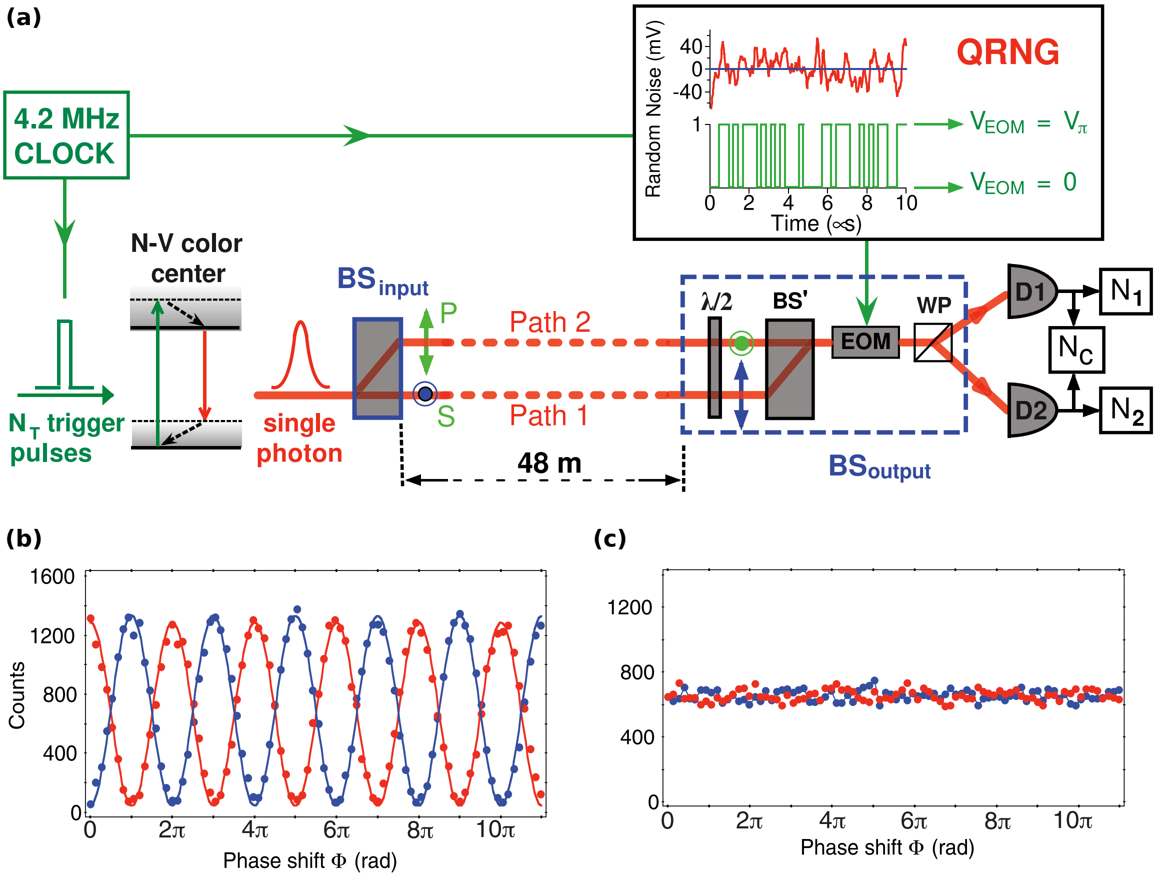


Figure 2.17: Experimental realization of Wheeler's delayed-choice experiment. Single-photons emitted by a single NV color center are sent through a 48-m polarization interferometer, equivalent to a time of flight of about 160 ns. A binary random number 0 or 1, generated by the QRNG, drives the EOM voltage between V = 0 and V = V_π within 40 ns after an electronic delay of 80 ns. The BS_{output} here comprises a half-wave plate, a BS′, an EOM, and a WP. The BS′ here is used to introduce phase shifts in the two paths. The voltage on the EOM controls the path information: Either no voltage is applied to the EOM, or its half-wave voltage V_π is applied. In the first case, the situation corresponds to the removal of the second BS (BS_{output}). In the second situation, the EOM is equivalent to a half-wave plate that rotates the input polarization by 45°, and this means inserting the BS_{output} in the path. Figure (b) shows the interference patterns of the two detectors when applying V_π to the EOM, and (c) is the count rate of the two detectors when applying no voltage to the EOM. Figures are adapted from [1].

number is created, nth photon enters the interferometer. Second step, the binary random number generated in the first step drives the EOM voltage to V =0 or V = V_π according to different bit values within rise time 40 ns and electronic delay 80 ns. In this figure, bit values for $n-1$th, nth, and $n+1$th photons are 1, 0, and 1. In the third step, the single-photon was recorded by detectors $D1$ and $D2$ after its flight time τ_{intef} in the interferometer. The detection was done during a gate of duration $\tau_{\text{d}} = 40$ ns. The blue shaded center zone in Fig. 2.18 represents the future light cone of the choice made by bit values. One can see that the photon entering the interferometer is clearly out of the future light cone associated with the random choice between the open and closed configurations.

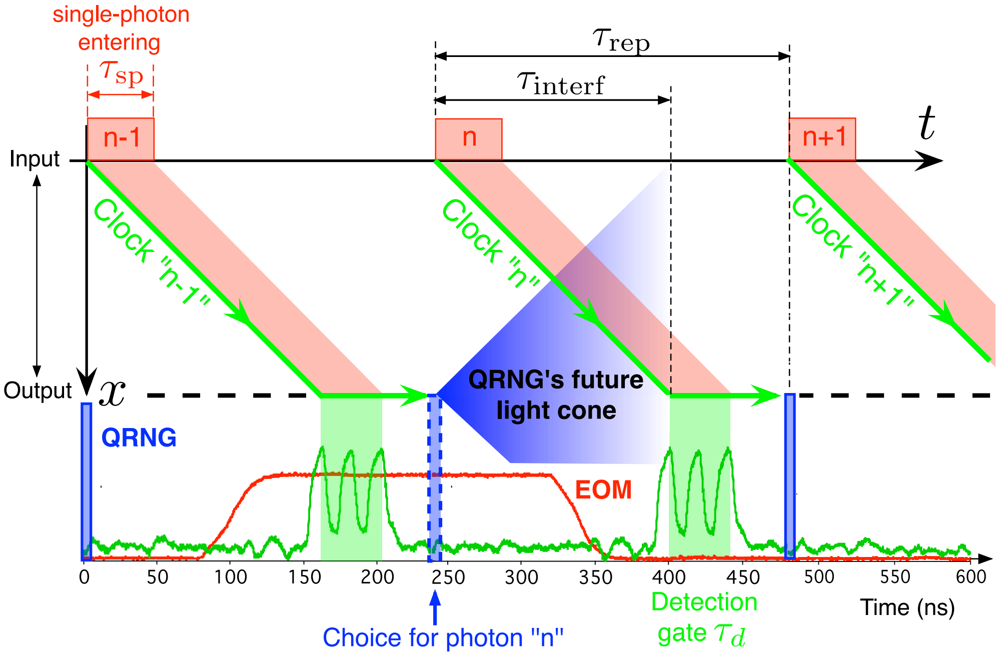


Figure 2.18: Timing of the delayed-choice experiment in [1]. This figure shows that the choice of whether to open or close the interferometer was spacelike-separated from the time when the photon entered into the interferometer. Figures are adapted from [1].

The realization of Wheeler's delayed-choice experiment demonstrates that nature behaves in agreement with the explanations of quantum mechanics.

2.2.4 Interference visibility

In the above delayed-choice experiment, the quantumness of the experimental data is related to the interference visibility of the interference pattern [132]. Similar to the fringe visibility [133], the interference visibility is defined as

$$v = \frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{max}} + I_{\text{min}}}, \tag{2.6}$$

where $I_{\max}$ and $I_{\min}$ are the maximum and minimum intensity in the interference pattern. In a Mach-Zehnder interferometer, $I_{\max}$ corresponds to maximum count rate $r_{\max}$ in the detectors, and $I_{\min}$ is the minimum count rate $r_{\min}$ of the detectors. so

$$v = \frac{r_{\max} - r_{\min}}{r_{\max} + r_{\min}}. \tag{2.7}$$

The interference visibility v is related to the count rate change of the detectors, and this change can only be caused by the interference of a quantum superposition state in the delayed choice experiment. So nonzero interference visibility means the photon is in a superposition state when the detectors detect it, and the knowledge of its path information is unknown [132]. When $v = 1$, it means photons are in pure quantum state $(|path1\rangle + |path2\rangle)/\sqrt{2}$ after they pass $\mathrm{BS}_{\mathrm{output}}$. When $0 < v < 1$, it means the click in the detectors are not only from the collapse of state $(|path1\rangle + |path2\rangle)/\sqrt{2}$, but also from some other sources including dark counts and external light sources. When $v = 0$, it means that the photons deterministically exist in path 1 and path 2 before they reach the detectors, and no interference happens.

2.2.5 From delayed-choice experiment to QRNG

The delayed-choice experiment can be easily turned into a QRNG. The delayed-choice QRNG, in principle, is a randomness expansion system where a private random seed is expanded into a long private random string. The private random seed is used to control the removal or insertion of the second beamsplitter in Fig. 2.13. The random seed is either bit 0 or 1. When it is bit 1, the second beamsplitter is inserted, the superposition state of the two paths will interfere, and an interference pattern will be observed in the two detectors. When the seed is 0, the second beamsplitter is removed, and the superposition state of the two paths will collapse into two detectors randomly and generate random numbers. The first beamsplitter creates the superposition state, but we do not need to trust it. Suppose there are samplers (shown in Fig. 2.10) after the first beamsplitter. In that case, the superposition state will be destroyed before it reaches the second beamsplitter to interfere, and the interference pattern will not be observed. So the visibility of the interference pattern is an indicator of the quality of the superposition of the two paths after the first beamsplitter. In our work, we use this interference visibility to quantify the quantumness of the experimental data. We connect this quantumness to the entropy of the generated raw bits by our randomness certification protocol in Chapter 5.

2.3 Bell test based QRNG

The Bell test was designed to test the statement of Bell's theorem, which is manifested as Bell's inequalities. The Bell test based QRNG is a device-independent QRNG. The advantage of this QRNG is obvious: true randomness can be generated without trusting the experimental devices since the quantum nature of the randomness in the experimental data is guaranteed by the violation of Bell's inequalities. This is a big difference compared to previous QRNG schemes: the single-photon QRNG requires the trust on beamsplitter and measurement devices, such as the detectors, and the delayed-choice QRNG puts trust in the measurement devices, including the detectors and the EOM. Before the illustration of the QRNG scheme, we first introduce Bell's theorem.

2.3.1 Bell's theorem

In 1935, Einstein, Podolsky, and Rosen (EPR) proposed a theory to point at the "inconsistencies" in quantum mechanics [134]. This theory was later referred to as the EPR paradox, which states that if one requires both realism and locality in a physical theory, it would lead to inconsistency in quantum mechanics. In such a theory, locality means that any signal, influence, or interaction propagates no faster than light. Realism means that one can assign properties to quantum systems before a measurement. The EPR theory opened the possibility of complementing quantum mechanics with local hidden variables (LHVs) to achieve realism. In the EPR theory, a bipartite entangled state is prepared. In this entangled state, if the position (or some other properties like polarization and spin directions) of the first particle was measured, then the result of measuring the position of the second particle can be predicted. Namely, EPR theory indicates that the properties of the entangled state are predetermined before the measurement.

In 1964, John Bell proved that quantum physics is incompatible with LHV theories [2]. Take two entangled photons as an example, and their polarization directions are entangled. In LHV theories, pre-existing values are the only local way to explain the perfect anti-correlations in the outcomes of polarization measurements along identical directions. But the pre-existing values are inconsistent with the predictions of quantum theory when the possibility of polarization measurements along different directions is allowed.

According to quantum theory, when polarization measurements along different directions are performed on the pair of particles in the entangled photons, the two opposite results (one “up” and one “down”) are not determined. Bell showed that the prediction of quantum mechanics for specific measurement scenarios differs from the predictions of all LHV theories [2]. Then Bell used inequalities (later referred to as Bell inequalities) to evaluate the validity of the EPR claims and any other local hidden variable theories. If Bell inequalities are violated, the quantum mechanic theory is correct; otherwise, the "inconsistencies" exist in quantum mechanics.

The requirements to do an experiment to violate Bell inequalities (such an experiment is called a Bell test) are very stringent. There are two loopholes [81] that need to be closed, the locality loophole [135] and the detection loophole [136]. Both loopholes in a Bell test were only closed simultaneously in less than a decade [61, 79, 80, 62].

2.3.2 CHSH inequality

The original Bell’s theorem did not clearly define the correlation function between different measurement settings and experimental results. In 1969, Clauser *et al.* proposed a different version of inequality with a well-defined correlation function between measurement settings and experimental results [73]. This equality is now called CHSH inequality. Several loophole-free Bell tests [61, 62] were based on the CHSH inequality because of its straightforward experimental scheme. The CHSH inequality was designed to guarantee that correlation between the two particles cannot be simulated classically as long as the inequality is violated.

In each round of the CHSH scenario Bell test, each party receives one particle from the entangled state and performs a local measurement on it using one out of two measurement settings. The measurements produce a binary output a for Alice and b for Bob. The choice of the local measurement settings depends on the randomly chosen binary input x for Alice and y for Bob.

The correlation value S of the CHSH inequality is

$$S = \sum_{x,y}(-1)^{xy}\left[P(a=b|xy) - P(a\neq b|xy)\right],$$

where $P(a = b|xy))$ (or $P(a \neq b|xy))$ is the probability that output $a = b$ (or $a \neq b$) when the measurement settings (x, y) are chosen. The above correlation equation can also be written as:

$$S = E(0,0) + E(0,1) + E(1,0) - E(1,1) \tag{2.8}$$

where $E(x, y) = P(a = b|xy) - P(a \neq b|xy)$. x, y have the same meaning as above. The upper bound of this correlation function in LHV theories is 2. Let the value of the "hidden variable" be λ, and it has a density function $\rho(\lambda)$. The integral of this density function over the entire hidden variable space is 1. With this hidden variable, we have:

$$E(x, y) = \int a(x, \lambda)b(y, \lambda)\rho(\lambda)d\lambda \tag{2.9}$$

where $a(x, \lambda)$ and $b(y, \lambda)$ are the measurement outcomes.

Then with four different measurement settings x, x', y, y',the following equations can be derived:

$$\begin{aligned}
&E(x,y) - E\left(x,y'\right) \\
=&\int \left[a(x,\lambda)b(y,\lambda) - a(x,\lambda)b\left(y',\lambda\right)\right]\rho(\lambda)d\lambda \\
=&\int [a(x,\lambda)b(y,\lambda) - a(x,\lambda)b\left(y',\lambda\right) \pm a(x,\lambda)b(y,\lambda)a\left(x',\lambda\right)b\left(y',\lambda\right) \mp \\
&a(x,\lambda)b(y,\lambda)a\left(x',\lambda\right)b\left(y',\lambda\right)]\rho(\lambda)d\lambda \\
=&\int a(x,\lambda)b(y,\lambda)\left[1 \pm a\left(x',\lambda\right)b\left(y',\lambda\right)\right]\rho(\lambda)d\lambda - \\
&\int a(x,\lambda)b\left(y',\lambda\right)\left[1 \pm a\left(x',\lambda\right)b(y,\lambda)\right]\rho(\lambda)d\lambda
\end{aligned}$$

By applying the triangle inequality, we have:

$$\begin{aligned}
&\left|E(x,y) - E\left(x,y'\right)\right| \\
\leq&\left|\int a(x,\lambda)b(y,\lambda)\left[1 \pm a\left(x',\lambda\right)b\left(y',\lambda\right)\right]\rho(\lambda)d\lambda\right| + \\
&\left|\int a(x,\lambda)b\left(y',\lambda\right)\left[1 \pm a\left(x',\lambda\right)b(y,\lambda)\right]\rho(\lambda)d\lambda\right|
\end{aligned}$$

Since both $[1 \pm a(x', \lambda) b(y', \lambda)] \rho(\lambda)$ and $[1 \pm a(x', \lambda) b(y, \lambda)] \rho(\lambda)$ are non-negative, so the left side equals to

$$\int |a(x, \lambda) b(y, \lambda)| \left|[1 \pm a(x', \lambda) b(y', \lambda)] \rho(\lambda) d\lambda\right| +$$
$$\int |a(x, \lambda) b(y', \lambda)| \left|[1 \pm a(x', \lambda) b(y, \lambda)] \rho(\lambda) d\lambda\right|,$$

which, when considering the fact that in the CHSH inequality, the outcomes are in binary format (in other words, $|a| \leq 1, |b| \leq 1$), is less than

$$\int [1 \pm a(x', \lambda) b(y', \lambda)] \rho(\lambda) d\lambda + \int [1 \pm a(x', \lambda) b(y, \lambda)] \rho(\lambda) d\lambda$$

Considering that the integral of $\rho(\lambda)$ over the complete hidden variable space is 1, that is $\int \rho(\lambda) d\lambda = 1$, we will have:

$$\begin{aligned} & |E(x, y) - E(x, y')| \\ & \leq 2 \pm \left[\int a(x', \lambda) b(y', \lambda) \rho(\lambda) d\lambda + \int a(x', \lambda) b(y, \lambda) \rho(\lambda) d\lambda\right] \\ & = 2 \pm [E(x', y') + E(x', y)] \end{aligned}$$

which means

$$\begin{aligned} & |E(x, y) - E(x, y')| \leq 2 + [E(x', y') + E(x', y)] \\ & |E(x, y) - E(x, y')| \leq 2 - [E(x', y') + E(x', y)] \end{aligned}$$

From this, it is straightforward to get

$$\begin{aligned} |S| &= |E(x, y) - E(x, y') + E(x', y') + E(x', y)| \\ &\leq |E(x, y) - E(x, y')| + |E(x', y') + E(x', y)| \leq 2 \end{aligned}$$

So from the above proof, we can see that within LHV theories, the upper bound of $|S|$ is 2, and the CHSH inequality is:

$$|S| \leq 2. \tag{2.10}$$

In contrast, quantum mechanics allows the value of S to be as large as $2\sqrt{2}$, known as Tsirelson's bound [73, 137]. Take one Bell state

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}} \tag{2.11}$$

as an example, this entangled state is composed of a pair of qubits. Alice has the first qubit, and Bob has the second one. Alice's two measurement observables are Pauli matrices σ_z and σ_x, the corresponding measurement setting are $x = 0$ and $x = 1$. Similarly for Bob, for two measurement settings $y = 0$ and $y = 1$, their observables are $-\frac{\sigma_1+\sigma_3}{\sqrt{2}}$ and $\frac{\sigma_1-\sigma_3}{\sqrt{2}}$. From the Born rule [57], the expectation values $E(x, y)$ for all four different combinations have the following values:

$$
\begin{aligned}
E(0,0) &= \langle \sigma_3 \otimes -\frac{\sigma_1+\sigma_3}{\sqrt{2}} \rangle = \frac{1}{\sqrt{2}}, E(0,1) = \langle \sigma_3 \otimes \frac{\sigma_1-\sigma_3}{\sqrt{2}} \rangle = \frac{1}{\sqrt{2}}, \\
E(1,0) &= \langle \sigma_1 \otimes -\frac{\sigma_1+\sigma_3}{\sqrt{2}} \rangle = \frac{1}{\sqrt{2}}, E(1,1) = \langle \sigma_1 \otimes \frac{\sigma_1-\sigma_3}{\sqrt{2}} \rangle = -\frac{1}{\sqrt{2}}.
\end{aligned}
$$

Then according to Eqn. 2.8, we have $|S| = 2\sqrt{2}$, which violates the CHSH inequality in Eqn. 2.10. Note that, for certain states and measurement settings combinations, a permutation of the values of a and b might be necessary to allow for a violation of the CHSH inequality

The experimental scheme of the first loophole-free Bell test is shown in Fig. 2.19, and the correlation value S of this Bell test violated the CHSH inequality, showing that nature is behaving in a quantum mechanic way.

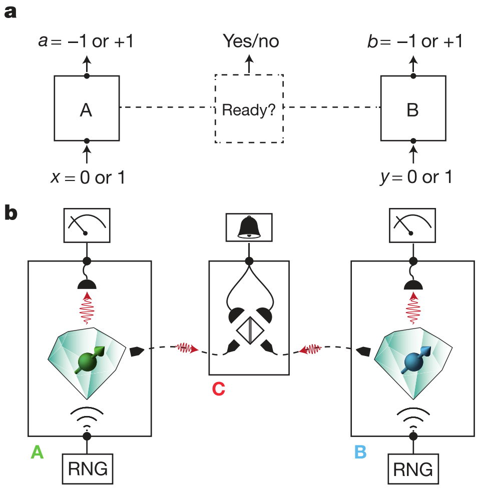


Figure 2.19: First loophole-free Bell test scheme. **a**, The bipartite Bell test scheme, two parties, Alice and Bob, accept binary input (x, y) to do corresponding measurements and produce binary outputs (a, b). This is done in an event-ready scenario, where an additional box between Alice and Bob gives a binary output signaling that the photons from Alice and Bob are successfully entangled. **b**, Experimental realization. The photons from Alice and Bob are generated from the NV center in diamond. A QRNG is used to provide the input (x, y). The electronic spin state in the NV is read out on a basis that depends on the input binary value, and the resultant signal provides the output. A box at location C records the arrival of single-photons that were emitted by Alice and Bob, and entangled with the spins of electrons at Alice and Bob's location. If the photons from Alice and Bob are successfully entangled at C, it means a bipartite entangled state between the electron spin states in Alice and Bob's NV center is prepared. Figures are adapted from [61].

2.3.3 QRNG scheme

The QRNG based on Bell test was first developed in 2009 [138], then a detailed randomness certification model and experiment were introduced and performed in 2010 [82]. Later, several loophole-free versions of this QRNG were presented [74, 75, 76, 77, 78].

The Bell test based QRNG is mainly based on the CHSH inequality. The quantum nature of the randomness in this QRNG is guaranteed by the violation of CHSH inequality, while the randomness itself is still from the random collapse of the superposition state. In the Bell test QRNG, the collapse of each entangled state involves two states at different locations. Take the Bell state $(|01\rangle + |10\rangle)/\sqrt{2}$ as an example. For this entangled state, if one measurement is performed in the first particle, the wave function of the second particle will also collapse simultaneously.

As long as the CHSH inequality is violated, there is quantumness in the experimental data, and this quantumness can be used to quantify the entropy of true randomness. The entropy can be quantified from the experimental data in different ways depending on different models. The details of the models will be discussed in Chapter 6.

2.4 Dimension witness based QRNG

A device-independent QRNG based on Bell's theorem has the highest security guarantee of the generated random number but is extremely challenging from the perspective of current technologies. On the other hand, the QRNG based on dimension witness offers a weaker form of security but can be implemented with standard technology, and it also has a relatively higher generation speed.

Before introducing the dimension witness, we emphasize the notations used here and in the corresponding chapters about dimension witness. x, y and z without hat "ˆ" are just labels, similar to any other label letters a, b, and their value depends on the context. When they have a hat ($\hat{x}, \hat{y}$ and $\hat{z}$), they represent different Bloch vectors in the corresponding axis of a Bloch sphere. When we want to refer to different axes in a Bloch sphere, we always use $x_{-\text{axis}}$, $y_{-\text{axis}}$, $z_{-\text{axis}}$ in this thesis. The Pauli matrices are represented as σ_1, σ_2, and σ_3 in this thesis. Using this notation system, our notations in this thesis are consistent with the ones in [83] and [4].

2.4.1 Dimension witness

The dimension witness is designed to distinguish the quantumness in the quantum systems, especially between classical and quantum systems. In quantum mechanics, experimental observations are associated with their Hilbert space. The dimension of Hilbert space can be decided by experimental data [139, 83]. This means we do not need to characterize all the experimental devices. A general formalism was developed to estimate the dimension of the classical and quantum system in a prepare-and-measure scenario. For the experiment scheme shown in Fig. 2.20, the dimension of the system is 2.

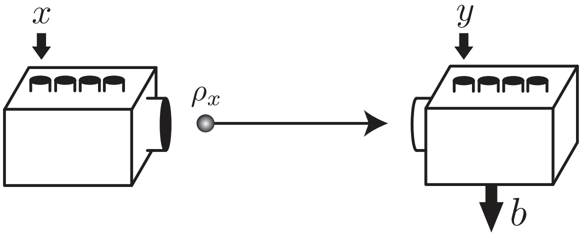


Figure 2.20: Prepare-and-measure scenario. The box on the left is the state source, it sends out state ρ_x to the box on the right side, which is the measurement device, and it performs measurement to the received state according to measurement setting y. The figure is adapted from [83]

The left device prepares a state ρ_x, and then the right device performs a measurement on the state. The observer tests the devices by choosing a preparation x and a measurement b, then receiving measurement outcome b. After multiple runs of this process, the observer obtains a probability distribution $p(b|x, y)$. The goal of the dimension witness is to provide a lower bound of the dimension of the states $\{\rho_x\}$ from the distribution $p(b|x, y)$ alone. The dimension witness here is nonlinear, and it is based on the determinant of a matrix [83]. Considering the simplest case, where $x = \{0, 1, 2, 3\}$, $y = \{0, 1\}$ and $b = \{0, 1\}$. This is a 2-dimension system, and the dimension witness for this system can be constructed as :

$$W_2 = \begin{vmatrix} p(0,0) - p(1,0) & p(2,0) - p(3,0) \\ p(0,1) - p(1,1) & p(2,1) - p(3,1) \end{vmatrix} \tag{2.12}$$

where $|.|$ means the determinant of this matrix, and $p(x, y)$ in the equation is $p(b = 0 \mid x, y)$. In general $p(b|x, y)$ can be described as:

$$p(b \mid x, y) = \rho_{\mathrm{x}} \mathrm{M}_{\mathrm{b|y}}, \tag{2.13}$$

where $M_{b|y}$ represents the measurement operators. Next, we show how large can W_2 be in a classical and a quantum system.

First, consider the situation of classical systems of dimension 2. According to the choice of preparation x, the left device sends out a classical message $m = 0, 1$. However, the left device has an internal source of randomness (this internal source can be anything unknown to the sender, including but not limited to the misalignment of the device and the shot noise.), which is λ_1, can affect the value of m. Thus the message m depends on both x and λ_1. The measurement device on the right side delivers an outcome b after receiving message m and measurement y from the observer. Similarly, the imperfection of the measurement device acts as an internal randomness source λ_2, which can affect the outcome b. So in the classical system, the probability distribution $p(b|x, y)$ is given by

$$p(b \mid x, y) = \int d\lambda_1 d\lambda_2 \rho\left(\lambda_1, \lambda_2\right) \sum_{m=0}^{1} p\left(m \mid x, \lambda_1\right) p\left(b \mid m, y, \lambda_2\right). \tag{2.14}$$

Suppose the state preparation device and measurement device are independent of each other, in other words, $\rho\left(\lambda_1, \lambda_2\right) = \rho_1\left(\lambda_1\right) \rho_2\left(\lambda_2\right)$. Then Eqn. 2.14 can be written as:

$$\begin{aligned} p(b \mid x, y) &= \int d\lambda_1 d\lambda_2 \rho\left(\lambda_1, \lambda_2\right) \sum_{m=0}^{1} p\left(m \mid x, \lambda_1\right) p\left(b \mid m, y, \lambda_2\right) \\ &= \sum_{m=0}^{1} \int d\lambda_1 \rho_1\left(\lambda_1\right) p\left(m \mid x, \lambda_1\right) \int d\lambda_2 \rho_2\left(\lambda_2\right) p\left(b \mid m, y, \lambda_2\right) \\ &= \sum_{m=0}^{1} s(m \mid x) t(b \mid m, y), \end{aligned} \tag{2.15}$$

where $s(m \mid x) = \int d\lambda_1 \rho_1\left(\lambda_1\right) p\left(m \mid x, \lambda_1\right)$, means the distribution of possible messages m for each state preparation x, and $t(b \mid m, y) = \int d\lambda_2 \rho_2\left(\lambda_2\right) p\left(b \mid m, y, \lambda_2\right)$ represents the distribution of outcomes b for measurement y when receiving message m. When choosing $b = 0$ as in Eqn. 2.12, $p(b \mid x, y)$ in Eqn. 2.15 can be written as $p(0 \mid x, y) = s(0 \mid x)[t(0 \mid 0, y) - t(0 \mid 1, y)] + t(0 \mid 1, y)$. Furthermore, the term $p(0|x, y) - p(0|x', y)$ can be represented as

$$p(0|x, y) - p(0|x', y) = S_{xx'} T_y \tag{2.16}$$

where $S_{xx'} = s(0 \mid x) - s(0 \mid x')$, $T_y = t(0 \mid 0, y) - t(0 \mid 1, y)$, and then we have

$$W_2 = \begin{vmatrix} S_{01}T_0 & S_{23}T_0 \\ S_{01}T_1 & S_{23}T_1 \end{vmatrix} = 0. \tag{2.17}$$

This suggests that for any strategy involving a classical bit, the determinant of W_2 is zero, which means the dimension witness value is 0.

Next, we discuss the performance of quantum strategies for the 2-dimension system in Fig. 2.20. This time, the source on the left side will send out state $\rho_x = (\mathbb{I}_2 + \vec{s}_x \cdot \vec{\sigma})/2$, and the measurement device will choose measurement operators $M_{0|y} = \left(c_y\mathbb{I}_2 + \vec{T}_y \cdot \vec{\sigma}\right)/2$, where $\vec{s}_x$ and $\vec{T}_y$ are Bloch vectors, and $|c_y| \leq 1$. Then we have

$$p(x, y) - p(x', y) = \operatorname{Tr}\left[(\rho_x - \rho_{x'})\, M_{0|y}\right] = (\vec{s}_x - \vec{s}_{x'})/2 \cdot \vec{T}_y. \tag{2.18}$$

Set $(\vec{s}_x - \vec{s}_{x'})/2 = \vec{S}_{xx'}$, then we get

$$W_2 = \begin{vmatrix} \vec{S}_{01} \cdot \vec{T}_0 & \vec{S}_{23} \cdot \vec{T}_0 \\ \vec{S}_{01} \cdot \vec{T}_1 & \vec{S}_{23} \cdot \vec{T}_1 \end{vmatrix} = \left(\vec{S}_{01} \times \vec{S}_{23}\right) \cdot \left(\vec{T}_0 \times \vec{T}_1\right) \leq 1. \tag{2.19}$$

The final inequality results from $\left|\vec{S}_{01} \times \vec{S}_{23}\right| \leq 1$ and $\left|\vec{T}_0 \times \vec{T}_1\right| \leq 1$. This bound for qubit strategies is tight and can be reached. For example the four preparation states can be chosen from $\vec{s}_0 = -\vec{s}_1 = \hat{z}, \vec{s}_2 = -\vec{s}_3 = \hat{x}$, where $\vec{x}$ and $\vec{z}$ are the unit Bloch vector. The Bloch vectors for the two measurement settings can be $\vec{T}_0 = \cos\theta\hat{z} + \sin\theta\hat{x}$ and $\vec{T}_1 = \sin\theta\hat{z} - \cos\theta\hat{x}$, where θ can be any value because of the rotational invariance of the cross product in the plane.

The above proof shows that only the qubit strategy can achieve $|W_2| > 0$. So, when we get $|W_2| > 0$, it suggests that the data involves the quantum measurement of a qubit. The definition of dimension witness in Eqn. 2.12 is also robust against noise. When there are effects of technical imperfections, including background noise and limited detection efficiency, a qubit strategy given by the data $p_Q(x, y)$ will achieve $|W_2| = Q > 0$. Suppose an error occurs with probability $1 - \eta$, for instance, one event is not detected, or some noise $p_N(x, y) = p_N(y)^2$ being detected, then we have $p(x, y) = \eta p_Q(x, y) + (1 - \eta)p_N(y)$, and the observed dimension witness is $W_2 = \eta^2 Q$, which is always positive when $Q > 0$. That

[2] The noise is independent of the choice of preparation x.

is to say, for an arbitrary amount of background noise or low detection efficiency, a qubit strategy will outperform any classical bit strategy.

In conclusion, the dimension witness defined in Eqn. 2.12 can demonstrate the advantages of using qubits in experimental systems, regardless of the imperfections of the devices. Since for any quantum strategies, we can achieve $|W_2| > 0$. For classical bits, only $W_2 = 0$ can be obtained. Note that there are two important preconditions of this dimension witness. The first is the independence between the state preparation device and the measurement device. Since if they are correlated with the classical bit strategy, one can reach $W_2 = 1$. For example, considering the equal mixture of the following deterministic strategies: $s(0|x) = 1$ iff $x = 0, 3$, and $t(0|m, y) = m + y$ mod 2, (ii) $s(0|x) = 1$ iff $x = 0, 2$, and $t(0|m, y) = m$. It is easy to verify that with this strategy, $W_2 = 1$ can be achieved. The second precondition is the limitation of the dimension. The W_2 is only valid for two dimension system. In higher dimension, with one classical trit, $W_2 = 1$ is also obtainable [83].

2.4.2 Randomness certified by dimension witness

The dimension witness introduced above can be used to certify quantum randomness since it can distinguish between classical and quantum systems. The randomness certified by dimension witness does not rely on a detailed model of the experimental devices, and this is relevant to real-world implementations. For example, it is well adapted to a scenario of trusted but error-prone device providers, i.e., a random number generator that is not actively designed to fool the user but where the implementation may be imperfect. The basic idea of this QRNG is to utilize the value of dimension witnesses, such as W_2 defined above [4]. When the value of W_2 is larger than zero, it means there is incompatible quantum measurement involved in the experiment results, and by Born's rule, this means true quantum randomness is generated. This true randomness can be directly quantified by the value of W_2 with a proper protocol, which is robust against other source of randomness, such as fluctuations from technical imperfections.

In Chapter 6 and 7, we discuss how to use the dimension witness defined by Eqn. 2.12 in different experimental schemes to certify true randomness. Especially in Chapter 6, we applied the dimension witness QRNG scheme for the first time to the Bell test data. We defined a remote state preparation dimension witness protocol that can certify randomness from Bell test data even when the CHSH inequality is not violated.

3 Methods and tools

The previous chapter explains the quantum phenomena used to generate random numbers in this thesis. When the raw random bits are generated, an important aspect to consider is how to quantify the entropy of randomness in the data. In this chapter, we illustrate how to quantify the entropy of randomness with min-entropy and how to extract uniformly distributed random sequences with the help of randomness extractors. After this, different test units in the NIST Statistical Test Suite are briefly introduced to show how this test suite qualifies for the structure of random sequences.

3.1 Min-entropy

Because of the deterministic nature of PRNGs, the PRNGs can be designed so that the random numbers from PRNGs are in a uniform distribution, and they can be put into use directly after their generation. In contrast, the random numbers from quantum processes cannot be guaranteed to be uniformly distributed because the true randomness contained in them cannot be easily tuned. For example, for the random numbers from radioactive decay, the random sequence is not uniform [27], and for photonic QRNGs built with biased beamsplitters. Besides, the randomness generated from quantum processes usually suffers from classical noises, including dark counts, thermal effects, and misalignments of the setup. All those classical noises do not contribute to quantum randomness. Thus, it is challenging to generate pure, uniformly distributed quantum randomness from a QRNG in practice. More often, the raw bits for a QRNG will be like

$$\begin{array}{c} 11101101010010001111111011111011010110100011010111 \\ 10011011110111110011111110011111110111100111011101 \end{array} \tag{3.1}$$

This sequence has more ones than zeros. However, a uniformly distributed sequence is usually preferred in the application of random sequences. A randomness extractor is needed to convert a biased random sequence to a uniformly distributed string. *Randomness extractors* are functions that can extract almost-uniform bits from raw and biased bits. The general structure of randomness extractors is shown in Fig. 3.1.

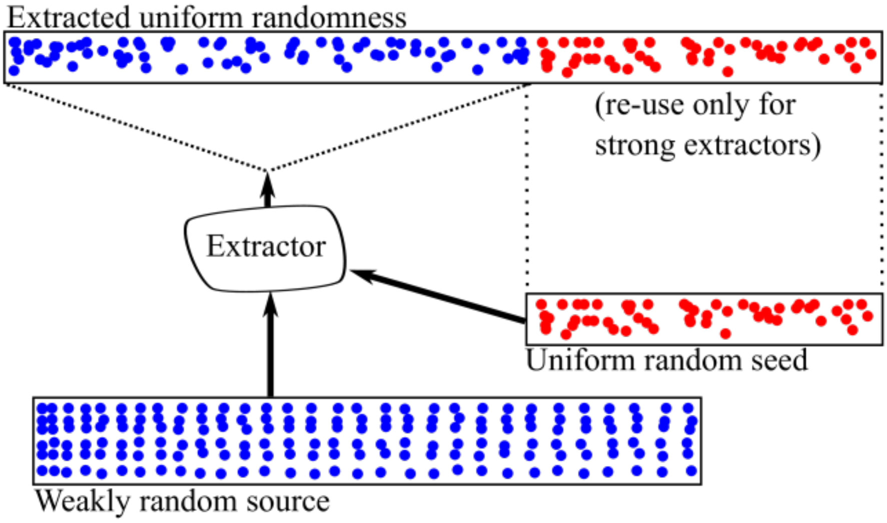

Extracted uniform randomness
(re-use only for strong extractors)
Extractor
Uniform random seed
Weakly random source

Figure 3.1: The structure of a seeded randomness extractor. Together with a short, uniformly distributed seed, the randomness extractor can make a weak, biased random source into a uniformly distributed strong random sequence. For strong randomness extractors, the initial seed is independent of the output and can be re-used, as shown in the figure. The figure is adapted from [140]

Mathematically, randomness extractor is defined as [141]

$$\mathrm{Ext} : \{0,1\}^n \rightarrow \{0,1\}^m, \tag{3.2}$$

where $\{0,1\}^n$ is the raw binary random bits with length n, and $\{0,1\}^m$ is a uniformly distributed binary random string. The extractors are originally designed to simulate random algorithms with weak classical sources. In recent years, extractors have had more applications in quantum cryptography and quantum randomness. More details of the randomness extractor are discussed in the next section.

In order to use randomness extractors, we need to know the entropy contains in the raw random bits from the experimental process. As mentioned in previous sections, entropy was first introduced by Shannon into information theory [21], where the entropy of a random variable is defined as the average unknown information inherent to the variable's possible outcomes. Shannon defined the entropy H (Greek letter eta) of a discrete random variable X with possible values $\{x_1, \cdots, x_n\}$ as:

$$\mathrm{H}(X) = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{3.3}$$

where p_i is the probability of x_i appearing in the random variable X.

In quantum random numbers, another entropy term–min-entropy–is often used instead of Shannon entropy. Min-entropy and Shannon entropy both belong to a general term called Rényi entropy. Rényi entropy is defined as [142]:

$$\mathrm{H}_\alpha(X) = \frac{1}{1-\alpha} \log \left(\sum_{i=1}^{n} p_i^\alpha \right) \tag{3.4}$$

where $\alpha \geq 0$ and $\alpha \neq 1$ is the order of the entropy. It is easy to verify that when $\alpha \to 1$, Eqn. 3.4 converges into Eqn. 3.3, which is exactly the definition of Shannon entropy. In the limit $\alpha \to \infty$, Eqn. 3.4 converges to the definition of min-entropy H_∞ (or $\mathrm{H}_{\mathrm{min}}$):

$$\mathrm{H}_\infty = -\log_2 \max p_i. \tag{3.5}$$

Compared with Shannon entropy, the min-entropy is more conservative. As shown in Fig. 3.2.

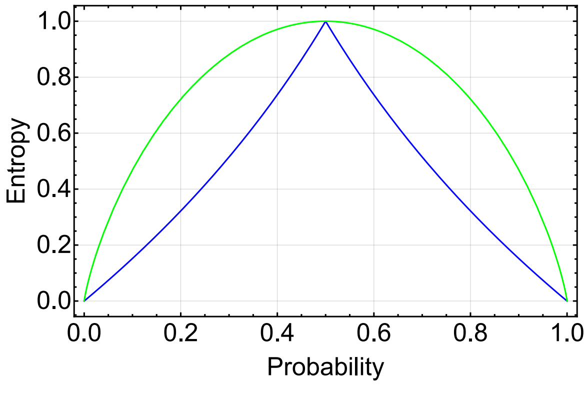


Figure 3.2: The comparison of Shannon entropy and min-entropy. The min-entropy is always not greater than the Shannon entropy

This figure shows that for a random sequence, its min-entropy is never greater than its Shannon entropy. The name “min”-entropy comes from the fact that it is the most conservative way to measure the information content of a discrete random variable in the family of Rényi entropies. For the biased sequence in Eqn. 3.1, it has 100 bits and its min-entropy per bit is 0.53 bits, so the total min-entropy of this sequence is about 53 bits, which means it contains the same amount of randomness as the following uniformly distributed 53 bits sequence

$$01100011111001001001110001011010110101100000001111010$$

The min-entropy of the sequence in Eqn. 3.1 is calculated directly from the sequence: The value $\max p_i = p_1 = 0.69$ because it has 69 ones. While for the raw bits from a QRNG, the $\max p_i$ should not be obtained from the raw bits, it must be estimated from quantumness in the raw bits to guarantee the generation of true randomness. The quantumness can be quantified differently in different quantum phenomena, which is discussed in Chapter 2.

The $\max p_i$ in a QRNG is also known as *guessing probability* p_{guess}, which represents the probability of correctly guessing the most probable value in one experiment run. The larger p_{guess}, the less random the source is, and the less min-entropy is there. In a QRNG, min-entropy is usually called conditional min-entropy. The word "conditional" comes from the fact that the p_{guess} is derived by conditioned on all the prior knowledge that *someone* has over the experimental device. This *someone* is often called an eavesdropper. The prior knowledge includes all the physical signals that can interact with the trust process. Let X be the random value from trusted physical processes and E be the prior knowledge known by the eavesdropper, and the conditional min-entropy can be defined as:

$$H_\infty(X|E) = -\log_2 \max p(X|E) \tag{3.6}$$

where $\max p(X|E) = p_{\text{guess}}$, which represents, with prior knowledge, the best chance the eavesdropper can correctly guess the next outcome value of the QRNG.

One of the main tasks in quantum random number generation is to build a model to estimate the conditional min-entropy for the corresponding QRNGs. The construction of the model depends on the level of trust of the QRNG device. Different trust levels assume different prior knowledge about the experimental devices. With a given level of trust, the QRNG model aims to maximize the conditional min-entropy in the generation process of the QRNG.

When the model is built and justified, the quantum randomness contained in the raw experimental data can be extracted with the help of randomness extractors.

3.2 Randomness extractor

There are different kinds of extractors [143, 144], and we now briefly introduce them. The first one is called deterministic extractors (assuming the distribution of the source is known). It is defined as

Definition 1 *Let $\mathcal{C}$ be a class of sources on $\{0,1\}^n$. An Δ-extractor for $\mathcal{C}$ is a function Ext : $\{0,1\}^n \to \{0,1\}^m$ such that for every $X \in \mathcal{C}, \mathrm{Ext}(X)$ is " Δ-close" to U_m,*

where $\{0,1\}^n$ can be treated as the raw bits with length n, and m is the length of the extracted random sequence. Δ can be understood as the error of the extractors, which represent the statistical distance between the output sequence $\{0,1\}^m$ and a perfectly uniformly distributed sequence U_m. Mathematically, statistical distance Δ is defined as

Definition 2 *For random variables X and Y from $\mathcal{U}$, their statistical difference (also known as variation distance) is $\Delta(X,Y) = \max_{T \subset \mathcal{U}} |\Pr[X \in T] - \Pr[Y \in T]|$. We say that X and Y are Δ-close if $\Delta(X,Y) \leq \Delta$.*

The deterministic extractors defined above are not applicable for unpredictable random sources, such as the raw bits from QRNGs, since the distribution of the raw bits is unknown. Even if the distribution of the raw bits is known, it will not be the same when the QRNG generates another sequence of raw bits. Thus, it is very inefficient if a randomness extractor has to be built for each run of the QRNG. For QRNGs, we need seeded extractors which can be used for indeterministic random sources. The definition of a seed randomness extractor is as follows:

Definition 3 *A function Ext: $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (k,Δ)-extractor if for every k-source X on $\{0,1\}^n$, $\mathrm{Ext}\,(X, U_d)$ is Δ-close to U_m.*

$\{0,1\}^d$ is a uniformly distributed random sequence with length d. It serves as the seed in the seeded extractors. The k-source X is defined as

Definition 4 *A random variable X is a k-source if $\mathrm{H}_\infty(X) \geq k$, i.e., if $\Pr[X = x] \leq 2^{-k}$.*

It is evident that the value k here is the min-entropy of the random variable (or random sequence) X, and $k \geq m$. This means the length of extracted bits m is smaller than the min-entropy k of the raw bits, and their difference is decided by the hashing error Δ of the extractors. The upper bound of Δ can be calculated by k and m [145]:

$$\Delta \leq \frac{1}{2}\sqrt{2^{m-k}}. \tag{3.7}$$

This equation means that the distance between $\{0,1\}^m$ and U_m is at most $\frac{1}{2}\sqrt{2^{m-k}}$.

There is one kind of seeded extractor called *strong extractor*. It is defined as

Definition 5 *Extractor Ext:* $\{0,1\}^n\times$ $\{0,1\}^d \rightarrow \{0,1\}^m$ *is a strong* (k,Δ)*-extractor if for every* k*-source* X *on* $\{0,1\}^n, (U_d, \mathrm{Ext}\,(X, U_d))$ *is* Δ*-close to* (U_d, U_m). *Equivalently,* $\mathrm{Ext}'(x,y) = (y, \mathrm{Ext}(x,y))$ *is a standard* (k,Δ)*-extractor.*

From this definition, we know that the random seed U_d in this extractor can also be used in the final uniformly distributed random sequence, which is shown in Fig. 3.1. Also, the extracted random sequence can be used as seeds for further randomness extraction for a strong randomness extractor. In this thesis, the extractors we used later are all strong extractors.

3.2.1 Toeplitz-hashing extractor

Two important strong randomness extractors can be used to extract a uniformly distributed random sequence from the experimental raw bits. These two extractors are Trevisan's extractor [146] and Toeplitz-hashing extractor [145, 147]. Trevisan's extractor has many theoretical advantages, including being secure against quantum adversaries and the seed length being polylogarithmic in the length of the input raw bits. However, the implementation of this extractor is more complex, and the output speed is much lower compared to Toeplitz-hashing extractor [147].

The Toeplitz-hashing extractor is specially designed universal hashing function — the Toeplitz hashing function, which is based on Toeplitz matrices [148]. There are two kinds of Toeplitz matrices: square and non-square. Let the square Toeplitz matrix be T_n, and the non-square one be $T_{n,m}$, then their definition is

Definition 6 *A square Toeplitz matrix $T_n = [t_{k,j}; k, j = 0, 1, \ldots, n-1]$ where $t_{k,j} = t_{k-j}$, i.e., a matrix of the form*

$$T_n = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & \cdots & t_{-(n-1)} \\ t_1 & t_0 & t_{-1} & & \\ t_2 & t_1 & t_0 & & \vdots \\ \vdots & & & \ddots & \\ t_{n-1} & & & \cdots & t_0 \end{bmatrix}$$

A non-square Toeplitz matrix $T_{n,m} = [t_{k,j}; k = 0, 1, \ldots, n-1; j = 0, 1, \ldots, m-1]$ where $t_{k,j} = t_{k-j}$, i.e., a matrix of the form

$$T_{n,m} = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & \cdots & t_{-(m-1)} \\ t_1 & t_0 & t_{-1} & & \\ t_2 & t_1 & t_0 & & \vdots \\ \vdots & & & \ddots & \\ t_{n-1} & & & \cdots & t_0 \end{bmatrix}$$

The above definition of the Toeplitz matrix shows that this matrix requires only the specification of the first row and the first column. The other elements are determined by descending diagonally from left to right. For the randomness extraction, a $T_{n,m}$ Toeplitz matrix is usually needed. Thus, the construction of the Toeplitz matrix needs a seed with $n + m - 1$ random bits, where n is decided by the length of the input raw bits and m is the total extractable entropy of randomness in the raw bits. The value of m can be calculated by a min-entropy value k in the corresponding QRNG model and a given error bound Δ (usually $\Delta = 0.001$) for the Toeplitz-hashing extractor. According to Eqn. 3.7, when we choose $m = k - 2\log_2 \frac{1}{2\Delta}$, the distance between $\{0, 1\}^m$ and U_m is at most Δ.

In general, with a given hashing error Δ, the implementation procedure of the Toeplitz hashing extractor is as follows:

- Suppose the size of the raw bits is n, the min-entropy of these raw bits is k and set the length of the output sequence m as $k - 2\log_2 \frac{1}{2\Delta}$.

- Prepare a random seed with $n + m - 1$ bits, and use it to construct the $T_{n,m}$ Toeplitz matrix.
- Multiply the raw bits ($1 \times n$) by the Toeplitz matrix ($n \times m$) to obtain a uniformly distributed random sequence of m bits, $\{0, 1\}^m$, which has a distance of at most Δ to U_m.

3.3 NIST Statistical Test Suite

Here we briefly introduce each test in the NIST [149] Statistical Test Suite. For the details of each test standard, please refer to [59, 149]. The NIST test suite is developed by NIST [150] to look at various aspects of a long random sequence. This test suite is a very important tool for understanding the structure of randomness. It has documented 15 statistical tests, and each statistical test is formulated to test a specific null hypothesis. The null hypothesis under each test is that the sequence being tested is random. The counterpart of this null hypothesis is called the *alternative hypothesis*, which states that the sequence is not random. For each applied test, a decision or conclusion is made to either accept or reject the null hypothesis.

For each statistical test included in the Test Suite, an appropriate randomness statistic must be selected and used to determine the acceptance or rejection of the null hypothesis. Assuming randomness, such a statistic has a distribution of possible values. Corresponding mathematical methods determine a theoretical reference distribution of this statistic under the null hypothesis. From this reference distribution, a critical value is derived. During a test, a test statistic value is calculated on the sequence being tested. This test statistic value is then compared to the critical value. If the test statistic value exceeds the critical value, the null hypothesis for randomness is rejected. Otherwise, the null hypothesis is not rejected. A P-value is usually used to summarize the strength of the evidence against the null hypothesis.

For all tests included in the NIST test suite, each P-value is the probability that a perfect random number generator produces a random sequence that is less random than the tested sequence. In other words, the larger the P-value, the more randomness in the tested sequence. To get a threshold to judge whether the test sequence is random or not under each test, a significance level α is usually chosen. If P-value$\geq \alpha$, the null hypothesis is accepted (the tested sequence is random). Otherwise, the null hypothesis is rejected. The value of α is usually 0.01, which indicates that one would expect one sequence in 100 sequences to be rejected. A P-value≥ 0.01 suggests that the sequence would be considered random with a confidence level of 99%, while a P-value< 0.01 means that the conclusion that the sequence is non-random has a confidence level of 99%.

The NIST Statistical Test Suite has 15 tests that were developed to test the distribution of given binary sequences. These 15 tests are:

(1) The Frequency (Mono-bit) Test,

(2) Frequency Test within a Block,

(3) The Runs Test,

(4) Tests for the Longest-Run-of-Ones in a Block,

(5) The Binary Matrix Rank Test,

(6) The Discrete Fourier Transform (Spectral) Test,

(7) The Non-overlapping Template Matching Test,

(8) The Overlapping Template Matching Test,

(9) Maurer's "Universal Statistica" Test,

(10) The Linear Complexity Test,

(11) The Serial Test,

(12) The Approximate Entropy Test,

(13) The Cumulative Sums (Cusums) Test,

(14) The Random Excursions Test,

(15) The Random Excursions Variant Test.

From broad theoretical considerations, the 15 tests above can be classified into four categories, namely Frequency Tests of ones and zeros (tests: 1 – 4), Tests for Repetitive Patterns (tests: 5 and 6), Tests for Pattern Matching (tests: 7 – 12) and Tests based on Random Walk (tests 13 – 15). Next, a high-level description of each particular test is presented.

(1) *Mono-bit Frequency Test.* This frequency test aims to test if the frequencies of ones and zeros across the entire sequence are close to $\frac{1}{2}$. If they are not equal to each other, it is intended to see if their difference falls in a tolerant error bound.

(2) *Frequency Test within a Block.* In the mono-bit frequency test alone, one can only test whether 0 and 1 appear with the same frequency. Their distribution is not taken into consideration. This test is intended to ensure that frequencies of 1s and 0s are evenly distributed across the entire sequence, so sequences like $10101010...101010$ would not pass the test.

(3) *Runs Test.* A *run* is an uninterrupted sequence of identical bits. In other words, a *run* of length k consists of exactly k identical bits, which are bounded by bits of opposite values. For example, 1100111001 has five *runs*. This test intends to see whether the number of runs of 1s and 0s of various lengths is as expected for a random sequence.

(4) *Longest Run Test of 1s in a Block.* The purpose of this test is to see if the frequencies of the longest run of 1s (or 0s) of various lengths appearing in the sequence are consistent with the length of the longest run of 1s that would be expected in a random sequence.

(5) *Binary Matrix Rank Test.* This test is intended to see if the tested sequence has repetitive patterns across its entire sequence. The tested sequence is sequentially divided into several disjoint blocks to see the linear dependence among its fixed length sub-sequence of each block. Each block is represented by a matrix of M rows and Q columns. Usually, both M and Q are taken as 32.

(6) *Discrete Fourier Transform (Spectral) Test.* This test checks if the tested sequence has periodic features across its entire sequence that indicate a deviation from the assumption of randomness. Considering randomness, one can find a peak height threshold value (T). If less than 5% of the peak heights are more than T, the tested sequence can be considered random.

(7) *Non-overlapping Template Matching Test.* This test aims to detect template matching in a non-overlapping manner, i.e., it looks for occurrences of pre-specified bit-sequence and to see if the numbers of such occurrences are within the statistical range of a sequence under the assumption of randomness.

(8) *Overlapping Template Matching Test.* Through this test, one can detect template matching in an overlapping manner, i.e., it looks for occurrences of pre-specified bit-sequence and

to see if the number of such occurrences is against a sequence under the assumption of randomness. Both this test and the previous non-overlapping Template Matching test using an m-bit window to search for a specific m-bit pattern. The difference between this test and the previous test is that: When the m-bit pattern is located, the window slides only one bit here, while in the previous test, the window is reset to the bit after the found pattern (in other words, the window slides m-bit).

(9) *Maurer's "Universal Statistical" Test* . This test aims to detect whether or not the tested sequence can be significantly compressed without loss of information. A random sequence is not compressible, and a significantly compressible sequence is considered to be a non-random sequence.

(10) *Linear Complexity Test.* The linear complexity test looks for the length of the Linear Feedback Shift Register (LFSR) and determines if the bit sequence from which the LFSR is obtained is random or not. The Berlekamp-Massey Algorithm is used here to obtain the LFSR. If the LFSR is long, the original bit sequence is considered random.

(11) *Serial Test.* If a sequence is random, then every m-bit pattern has the same probability of appearing as every other m-bit pattern. The serial test counts the frequency of all possible overlapping m-bit patterns across the entire sequence. One can see if the sequence can be termed as random or not based on the deviations of each of all the counts together,

(12) *Approximate Entropy Test.* This test is a test of randomness based on repeating patterns. The larger the entropy, the more randomness in the tested sequence. For the tested sequence, the approximate entropy is measured by comparing the frequency of overlapping patterns of all possible m-bit patterns with that of $m + 1$-bit patterns. If the approximated entropy is smaller than a threshold value, the tested sequence is considered non-random.

(13) *Cumulative Sums Test.* In this test, the bits (b_i) in the tested sequence are converted to X_i (-1 or $+1$) using $X_i = 2b_i - 1$, and the cumulative sum S_i is a series of sum of the converted values X_i: $S_1 = X_1,\ S_2 = X_1 + X_2, ..., S_n = \sum_{i=1}^{i=n} X_i$. This test determines whether the cumulative sum of the partial sequences S_i in the tested sequence deviates from the expected behavior of such a cumulative sum for random sequences. This cumulative sum can be considered a random walk. For uniformly distributed random sequences, the excursions of the random walk are near zero.

(14) *Random Excursions Test.* The Random Excursions Test checks whether the number of cycles with exactly k visits to a given state in a cumulative sum random walk deviates from what is expected in a random sequence. For example, for a cumulative sum $S = \{0, -1, 0, 1, 0, 1, 2, 1, 2, 1, 1, 0\}$, there are three cycles:$\{0, -1, 0\}$, $\{0, 1, 0\}$, and $\{0, 1, 2, 1, 2, 1, 1, 0\}$. This test examines if the number of visits $k (= 0, 1, ..., 5)$ to a particular state in one cycle deviates from the visit in a random sequence. The *state* here means the value in each cycle, like -1 in the cycle $\{0, -1, 0\}$. Continuing with the example, when $k = 0$ (means no visit), the "no visits" to state -1 among all three cycles is 2 since -1 does not appear in two cycles. Similarly, when $k = 1$ (means visit only once), the "one visit" to state -1 is 1 since -1 only appears once in all three cycles.

(15) *Random Excursions Variant Test.* This test looks for the total amount of visits to a particular state in cumulative sums of a random walk across the entire bit sequence. It detects the deviations from the expected number of visits in the random walk of a random sequence. The difference between this test and the previous test is that the number of visits in this test is directly counted in the given cumulative sum without considering the cycles. Take the same cumulative sum $S = \{0, -1, 0, 1, 0, 1, 2, 1, 2, 1, 1, 0\}$ as an example, the total number of visits to state -1 is 1, to state 1 is 5, and to state 2 is 2.

4 Randomness from a single-photon source based on NV center

A variety of QRNGs have been implemented at the beginning of this century based on the measurement of photonic qubits [151, 94]. Later, with the increasing demand for speed and the implementation of coherent states, several generators have been implemented using light sources other than single-photons. These generators often measure the vacuum fluctuations [152, 153, 154] or phase noise [155, 156]. Here we propose a photonic QRNG based on a NV single-photon source and use the antibunching characteristic of single-photons to strengthen the randomness in the experimental data. The basic principle is discussed in Chapter 2. This chapter focuses on the experimental setup, model construction, and data analysis of our proposed photonic QRNG. More details of this work can be found in our publication [157].

4.1 Experimental setup

The basic scheme of the experiment is shown in Fig. 4.1(a). A fiber guides a stream of single-photons toward a beamsplitter. This beamsplitter's second input arm is blocked, commonly described as a vacuum state ($|0\rangle$). Therefore, the single-photons are distributed on the beamsplitter according to the beamsplitter ratio. The photons are detected on the detector A (transmitted photons) and detector B (reflected photons). This setup has been operated for seven days in continuous operation.

In the experiment, the NV-based single-photon source is optically excited by a continuous wave laser, and the resulting fluorescence is detected by confocal microscopy (as shown in Fig. 4.1a). The experiment is operated under an ambient condition and spans less than 1 m^2 of an optical table.

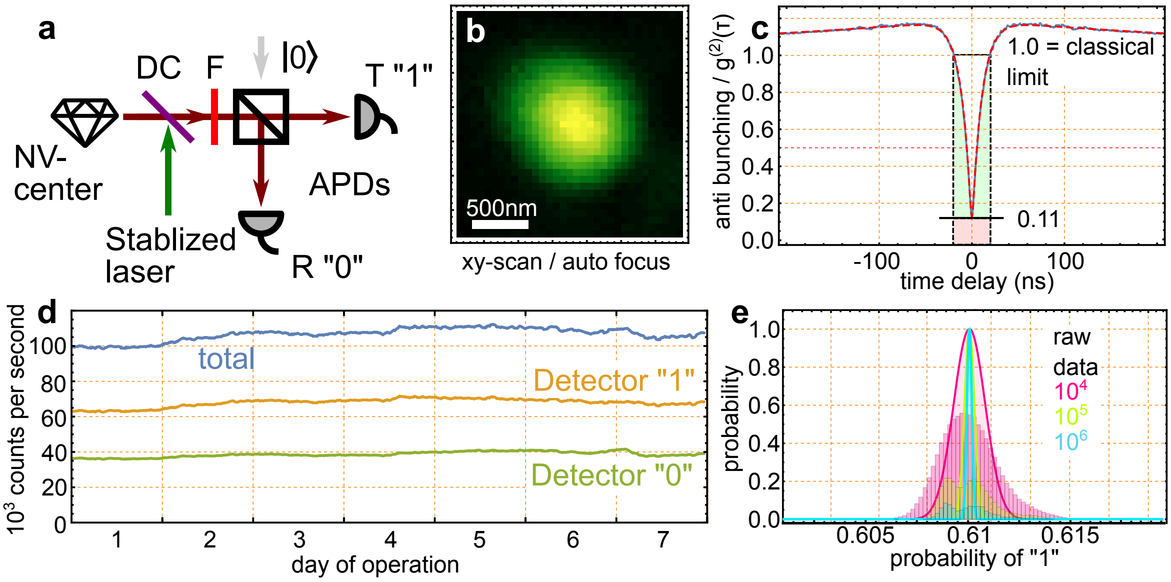


Figure 4.1: Experimental Configuration. **a**, The scheme of the single-photon random number generator. A confocal microscope is used to locate a single NV center. A single-photon is measured with two avalanche photodiodes (APDs). DC: Dichroic Mirror; F: Long-pass Filter. **b**, Fluorescence counts of a lateral scan over the diamond sample. Peak intensity is about 100 kcps (kilo counts per second). **c**, Measurement of antibunching and a theoretical fit (dashed line in the figure), the timing resolution here is 0.5 ns. **d**, A long time recording of the raw bits, the exact time is 608125 seconds. **e**, The raw data is biased due to the unbalanced beamsplitter in the setup. The figures are adapted from our publication [157].

The laser (with wavelength λ=532 nm), which is used to excite the single emitter inside the NV center, is operated in a continuous wave mode. To avoid laser power fluctuations, the intensity is stabilized by a commercial PID controller (Stanford Research, SIM960). For this, the laser power is measured shortly before the diamond single-photon source. An acoustic-optical modulator regulates the laser power at the laser output.

After the laser beam is reflected off a dichroic mirror, it is guided to a galvanometric mirror system, and then the laser beam is steered into a $4f$-scanning microscope. The focus is realized by a 100× oil objective (Olympus Plan FL N, NA=1.35). In the confocal configuration, the emitted single-photons are captured by the same microscope objective, guided backward, and transmits through the dichroic mirror toward the detection system. To suppress the stray light, the detected light is then focused (f=100 mm) onto a pinhole ($\varnothing$=50 μm) and filtered by a 640 nm long pass filter. The detected light is then transferred by $2f$-$2f$ imaging through a symmetric non-polarizing beamsplitter towards two single-photon detectors (Count, Laser Components). This configuration reduces the avalanche photodiode (APD) cross-talk significantly.

The sample is a mm-sized diamond that hosts nitrogen-vacancy centers at natural abundance. For high excitation and collection efficiency of the NV center, a solid-immersion lens was fabricated around an earlier confocally localized center [112]. The NV centers and the solid-immersion lenses are identified by confocal beam scanning. The scanning beam was used not only to locate the NV centers but is repeatedly (Δt=8 min.) medially and laterally scanned across a certain area during the experiment. After which, the NV center is re-centered, and the measurement is continued. This compensates for the sample's drift during the experiment run. One of the lateral images is shown in Fig. 4.1b.

All detection events in the two APDs are recorded on an FPGA-based time tagger (Swabian Instruments, Timetagger 20). The time resolution of the time-tagger is 100 ps. Each click event is recorded in a 128 Bit binary format (64-bit, which detector has clicked, and 64-bit with the time in ps). In the 7 days experimental time (608125 seconds, including the refocusing periods), we recorded 832 GiB raw bits, and the average count rate is about 91.7 kcps. The experimental time without refocusing is approximately 558,000 s, and the average count rate, in this case, is around 100 kcps. The total 832 GiB raw bits set are split into 179 files, which are analyzed below.

To prove the single-photon nature of the emitted photon stream, the antibunching effect of the photons is analyzed with a HBT configuration in Fig. 4.1a (see also Fig. 2.5). This is performed by correlating the recorded time stamps of the two APDs in a start-multiple-stop fashion [158]. The corresponding antibunching curve is shown in Fig. 4.1e. It shows an antibunching "dip" below the value of $g^{(2)}(0)$=0.5, which proves the single-photon nature of the source. The timing resolution τ_{rs} for the start-stop event is 500 ps. The bunching behavior in the curve is due to the NV centers' typical meta-stable state between the two level states, as shown in Fig. 2.9.

During the experiment run, the whole setup was covered with blackout material without human interaction. A measurement of the count rates is shown in Fig. 4.1d. In the course of the raw bits recording, some fluctuations are observed. These are caused mainly by the thermal drift of the table, which affects the position of the pinhole and both APDs.

4.2 Bound the entropy in the raw bits

The outcomes from two discrete single-photon detectors in different output modes of a beamsplitter are interpreted as single raw bits. These bits then go into the randomness extraction process. To extract randomness, we first need models to quantify the entropy in the raw bits. We introduce three models to quantify the entropy in the raw bits from single-photons.

4.2.1 Randomness generation from raw bits

In the first model, we simply consider *all* click events in the APDs as raw random bits. The probability of the individual outcome and the transition probabilities are relevant. Our interpretation here in the first model is not limited to using a single-photon input state.

The probability of whether a photon is reflected or transmitted in a beamspliter is linked to the vacuum state at the second input port of the beamsplitter. Further, the probability also depends on a variety of experimental factors. Most importantly, the beamsplitter ratio $\mathcal{R}$ is a function of the reflection (R) and transmission (T) coefficient. Our model assumes a loss-less beamsplitter, i.e., $T + R = 1$. Although a biased beamsplitter introduces an imbalance in the raw bits, it does not introduce any memory in the experimental setup.

Another crucial parameter is the detector efficiency of the utilized detectors, $\eta_{\{\mathrm{A,B}\}}$. This value describes how many incident photons lead to an electrical pulse that can then be recorded. The detector efficiency of each detector is closely linked to the beamsplitter ratio in a given experimental implementation. An electrical pulse from a single-photon detector has a finite length, so it can be recorded with normal detection hardware. After one electrical pulse is generated from the detector, a second detection event is usually suppressed. This suppression time is called the “dead-time”, τ_{dead}, of the detector. In the common Geiger mode photodetector modules, the time when no second photon is detected usually exceeds the length of the electrical pulse.

The technicality of the detector’s dead times also introduces another problem for generating raw random bits: two subsequent clicks from two detectors can be correlated since one of the detectors is in its dead time, and only the second detector is capable of detecting another incoming photon. This is not the focus of our work; for a detailed study of this situation, please refer to [158].

With given beamsplitter ratio and detector dead time, the total count rate of the two detectors is $r_{\mathrm{total}} = r_{\mathrm{A}} + r_{\mathrm{B}}$. It amounts to the following expression,

$$\begin{aligned} r_{\mathrm{A}} &= \eta_{\mathrm{A}} T I_{\mathrm{in}} - \frac{(\eta_{\mathrm{A}} T I_{\mathrm{in}})^2 \int_0^{\tau_{\mathrm{dead}}^{\mathrm{A}}} g^{(2)}(\tau) d\tau}{4}, \\ r_{\mathrm{B}} &= \eta_{\mathrm{B}} R I_{\mathrm{in}} - \frac{(\eta_{\mathrm{B}} R I_{\mathrm{in}})^2 \int_0^{\tau_{\mathrm{dead}}^{\mathrm{B}}} g^{(2)}(\tau) d\tau}{4}, \end{aligned} \tag{4.1}$$

where $\eta_{\mathrm{A,B}} T I_{\mathrm{in}}$ is the click rate of the detector when there is no dead-time and $g^{(2)}(\tau)$ is the antibunching curve [159] of the utilized single-photon source.

This equation represents the click rates of the two detectors independently of the input source. At the same incident photon flux, a single-photon source has a higher probability of a later detection event than a laser with the same brightness since a laser source obeys an exponential decaying probability distribution in the subsequent detection of a photon. This is related to the detector's dead time and the count rates, as outlined in more detail in [158].

The detection rates for the two detectors have been determined by considering the beamsplitter ratio, the detection efficiency, and the dead time. The *probabilities* for calculating the min-entropy below can be experimentally determined straightforwardly by the ratio of the detector events:

$$p_{\mathrm{A}} = \frac{r_{\mathrm{A}}}{r_{\mathrm{A}} + r_{\mathrm{B}}} \tag{4.2}$$

The details of p_{A} and r_{A} can be found in the appendix. The conditional probabilities are calculated as follows. As an example, we only show the derivation for the case of the conditional detection of $p(A|A)$:

$$p(A|A) = \big(1 - \int_0^{\tau_{\mathrm{dead}}^{\mathrm{A}}} g^{(2)}(\tau) d\tau\big) \eta_{\mathrm{A}} T \tag{4.3}$$

This conditional probability is affected by the combination of the detector efficiencies and the beamsplitter ratio.

In this model, the entropy in the raw bits relates to the bias and other technical considerations, such as the detector efficiency and the dead time of the detectors. The calculated and conditional probabilities are utilized to calculate the conditional min-entropy.

The conditional min-entropy is used in post-processing the raw random bits. The definition of the conditional min-entropy, H_∞ is given as [160][1]:

$$H_\infty(X|Y) = -\mathrm{log}_2\Big(\sum_y p(y)\max_x\{p(x|y)\}\Big)\,, \tag{4.4}$$

where x and y are two subsequent events in the raw random bits. In our case, $\{X, Y\} \in \{0, 1\}$, subsequently the conditional min-entropy is

$$\begin{aligned} H_\infty(X|Y) &= -\log_2\left(\sum_y p(y)\max_x\{p(x|y)\}\right) \\ &= -\log_2(p(0)\max\{p(0|0), p(1|0)\} + p(1)\max\{p(0|1), p(1|1)\}) \end{aligned} \tag{4.5}$$

After the min-entropy of the raw bits is quantified, uniformly distributed random sequences can be extracted by randomness extractors. More specifically, the value of $H_\infty(X|Y)$ can be quantified using the relevant probabilities derived from the experimental parameters; see the appendix for details. Suppose $H_\infty(X|Y) = k$, since the generation speed of the raw bits is r_{total}, then the generation speed of the uniformly distributed random bits amounts to kr_{total}.

4.2.2 Randomness generation from single-photon events

The above model can also be applied to a random number generator with a simple laser or even a light-emitting diode (LED) input source. However, using a laser has drawbacks: When a coherent state $|\alpha\rangle$ is present, the user has no clue if the device *really* detects the incoming (laser) mode. As shown in Fig. 1.6, the two beams imping from a laser via the beamsplitter onto the detectors are unrelated in several ways. They cannot tell if two different laser sources (or two independent laser modes) are observed with different detectors. In

[1] Note that the definition of conditional min-entropy here is equivalent to the definition in Eqn. 3.6. In this case, we suppose the previous event y is known (belongs to the knowledge of E), then we derive the guessing probability of the current event x

the worst case, an eavesdropper can remotely control the detector clicks, and due to their uncorrelated origin, the device owner has no proof of their origin or integrity.

A single-photon source is advantageous at this point and can enhance the trust in the random bit generation scheme. An auto-correlation function can be recorded to prove the existence of a single-photon emitter. For a single-photon source, this shows the typical single-photon antibunching, which allows us to characterize the non-classicality of the source, and it also gives an upper bound on the amount of (uncorrelated) counts, which an eavesdropper might control. The advantage of a true single-photon source is that it allows a user to exclude several attack scenarios and to guarantee the independence of an external adversary of the device to a certain extent.

Antibunching is commonly described by the auto-correlation function $g^{(2)}(\tau)$, and it cannot be interpreted by classical theory [159]. All the spurious background contributions introduced are uncorrelated events that can be known to an external adversary, and they will change the antibunching curve of a single-photon emitter and increase the value of $g^{(2)}(0)$.

The term *true single-photons* denotes all the photons detected by the detectors and stem from the device-internal single-photon source. Usually, the non-ideal measurement devices reduce this fraction from 100%. This fraction can be determined by the $g^{(2)}(0)$ value as $\sqrt{1 - g^{(2)}(0)}$ [111]. This allows us to estimate the "quantumness" of the single-photon stream.

When $g^{(2)}(0) \geq 1$, all the raw experimental data will be discarded. In this case, the source for the random number generator is no longer based on the device's internal single-photon source [159, 161]. The random number generation process is likely to be externally affected and, in the worst case, completely controlled by an eavesdropper. When $0 \leq g^{(2)}(0) < 1$, there is "quantumness" involved in the click events, and we can bind this "quantumness" to the amount of randomness generation (under fair-sampling assumption).

In order to quantify the entropy in the raw bits, the calculation of the guessing probabilities is needed. For this, we refer to the discussion above. Considering background events implies that a few raw bits are known to an eavesdropper. Therefore it is not clear if the genuine device generated them or if they were intentionally introduced in an uncorrelated manner. This fraction is accounted for by introducing a probability p_e, which means that the eavesdropper knows the value of the raw bits in this fraction.

We now calculate the min-entropy of the stream with the unwanted background fraction. This requires that our model be extended with a third possible outcome associated with an eavesdropper or classical background light. The click events are represented as A, B, and the events known to an eavesdropper are represented as E. The fraction p_e of E reduces the overall amount of entropy of the random number generator. The fraction of pure single-photon events is $\sqrt{1 - g^{(2)}_{\text{fit}}(0)}$. Then p_e can be calculated as: $p_e = 1 - \sqrt{1 - g^{(2)}_{\text{fit}}(0)}$.

Now, the min-entropy can be calculated with the following equation:

$$\begin{aligned} &H_\infty(X|Y) \\ &= -\log_2\Big(p_e + (1 - p_e)\big(\sum_y p(y) \max_x\{p(x|y)\}\big)\Big) \,. \end{aligned} \tag{4.6}$$

Compared to the above Eqn. (4.5), the extractable entropy given by the Eqn. (4.6) is reduced. However, this equation can quantify the random bits generated by single-photon events, meaning they originate from the generator's genuine source. Suppose $H_\infty(X|Y) = k_{\text{q}}$, then with an n bits long raw random bits, $k_{\text{q}}n$ true random bits can be generated. The generation speed of true randomness corresponds to $k_{\text{q}}r_{\text{total}}$.

4.2.3 Conditioned tuple detection of detector clicks

In the above model, we assume that the recording of the antibunching curve describes the whole data stream, even if a single-photon was detected with no further closely neighboring detection events. This means that the photons not contributing to the small time window of the antibunching dip (approx. 1-30 ns, corresponding to the T$_1$-time of the system) are considered to be single-photons from a legitimate source. Ideally, a normalized antibunching curve is below unity, but in practice, because of the thermal fluctuations and some other imperfections of the experimental devices, the antibunching exceeds the value of unity. To further guarantee the quantumness in the random bits, we now only consider the clicks in a time window below the line of unity in the antibunching curve. This implies a time-wise selection of events with a "partner-photon" temporally close by. In other words, only the *tuples* of photon detection events are considered. In the area below the antibunching curve, we take the tuple event "AB" as one random bit and "BA" as the other random bit (we are dealing with binary random bits, so we only have two random bits: 0 or 1.).

Tuple events are often described as “start-stop events” on two single-photon detectors behind a beamsplitter. In principle, tuple events “AB” and “BA” are balanced since their detection probabilities are equal $p(\mathrm{AB}) = p(\mathrm{BA})$, as long as the experimental parameters do not change in the course of the experiment. By considering the tuple events only, the detection rate of the raw randomness events reduces drastically since only start-stop events in a limited time range are valid. These events are anti-correlated but are still above the background noise level in the $g^{(2)}(\tau)$-recording.

Ideally, a normalized antibunching curve of an ideal two-level system single-photon source exists only below unity. This implies that the raw events all fulfill the non-classical nature of $g^{(2)}(\tau) < 1$. In a real-world experiment, some parameters of the experimental devices may have fluctuations. These may be introduced as shot noise. To guarantee the quantum nature of the utilized single-photon source, in this situation, we consider the data below unity with a given standard deviation, for example, 11.5σ, which significantly limits the probability of an outlier.

An external adversary now has limited options to influence the device based on the strategy of tuple detection: Any uncorrelated event that is controlled by an external eavesdropper will lead to an increased background fraction. Therefore, an external eavesdropper would have to implement more sophisticated strategies to launch clicks in the generator. For example, when the primary process launches a click, i.e., an emitted photon from the single-photon source, the eavesdropper has to launch a click onto the other detector of the generator within the dead time of the first detector. This requires a stringent timing of the clicks launched by the eavesdropper. First, the eavesdropper has to detect that there has been a click in the generator, and then she has to *introduce another click* into the generator within a very short legitimate time. This requires that the eavesdropper be very close to the generator due to the signal traveling time. In this sense, based on short time differences (ns) in subsequent clicks in the generator, the generated bits are “fresh” and guaranteed to be unaltered for a short time.

Another, more sophisticated attack version from the eavesdropper can be to fake the clicks on both detectors as if this eavesdropper had another single-photon source at her location. Then, the number of clicks outside the small nanosecond range of the primary source is large. Depending on the relative brightness compared to the primary source, this will lead to a certain amount of background clicks again. The eavesdropper’s only option is to fully suppress the primary emission of the single-photon source and replay an equivalent detector control scheme which would introduce a comparable antibunching signal. But a simple control of the primary photon source can reveal that an external adversary is active.

To calculate the probabilities to bound the min-entropy, the tuple event probabilities $p(\mathrm{AB})$ and $p(\mathrm{BA})$ are considered. Furthermore, the conditional probabilities (e.g. $p(\mathrm{AB}|\mathrm{AB})$) have to be described. For a full derivation, see the appendix.

In this model, only paired events and the *area* below unit line (i.e. $g^{(2)}_{\mathrm{fit}}(\tau) <= 1$) are considered. This area is shown as green in Fig. 4.2. This unit line is the classical limit [159, 161, 109]. This area determines the fraction of single-photon quantum randomness from the raw data in a very conservative way. To estimate this fraction, the click rates of the raw events are required. For the convenience of description, this area is named the "quantum area" in the following.

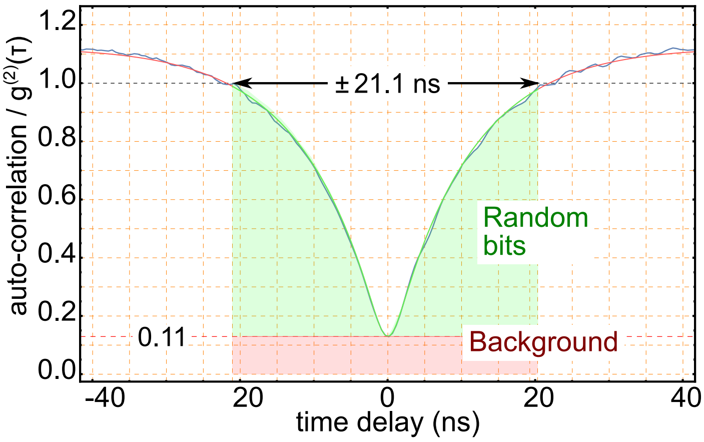


Figure 4.2: Antibunching as a measure for quantumness. The antibunching effect of single-photons is only observed in a small time window. In our third randomness extraction model, the area of the generated bits between the classical bound of $g^{(2)}(\tau) \leq 1.0$ and above the background level are considered. This reduces the number of raw input bits for the generator dramatically. The figures are from our publication [157].

The quantum randomness fraction can be derived by considering the tuple events for a given timing resolution τ_{rs}. The tuple event rate r_{stsp} of the two detectors and uncorrelated events (e.g. laser emission) is given for a certain time resolution τ_{rs} as:

$$r_{\mathrm{stsp}} = r_{\mathrm{A}} \times r_{\mathrm{B}} \times \tau_{\mathrm{rs}} \, .$$

In the "quantum area", at different delay times, the tuple events within a given timing resolution τ_{rs} correspond to different $g^{(2)}(\tau)$ values. This means they obey different probabilities [109]. The experimental antibunching curve is represented as $g^{(2)}_{\mathrm{fit}}(\tau)$. Then, the total photon tuple event rate in this quantum area is given by

$$r_{\mathrm{A}} \times r_{\mathrm{B}} \times \sum_{\tau=-t}^{\tau=t} \tau_{\mathrm{rs}} g^{(2)}_{\mathrm{fit}}(\tau) \approx r_{\mathrm{A}} \times r_{\mathrm{B}} \times \int_{-t}^{t} g^{(2)}_{\mathrm{fit}}(\tau) d\tau \,,$$

where t satisfies $g^{(2)}_{\mathrm{fit}}(t) = 1$, which means that the whole range is considered until the events are not anti-correlated anymore. $g^{(2)}_{\mathrm{fit}}(\tau)$ is the antibunching curve with background noise, which means the above equation also contains the start-stop events caused (partially) by uncorrelated background noise. We use the fraction $\sqrt{1 - g^{(2)}_{\mathrm{fit}}(0)}$ to discard the background noise in the clicks of each detector. Then the tuple events originating from single-photon events can be calculated as

$$r_{\mathrm{stsp}} = (1 - g^{(2)}_{\mathrm{fit}}(0)) \times r_{\mathrm{A}} \times r_{\mathrm{B}} \times \int_{-t}^{t} g^{(2)}_{\mathrm{fit}}(\tau) d\tau \,. \tag{4.7}$$

The count rate in this area is the generation speed of single-photon events, which are short-time related. Therefore, the generation speed of the true randomness in this part is linked to the tuple events as $r_{\mathrm{rand}} = r_{\mathrm{stsp}}$. Increasing the excitation laser power will increase the single photon generation speed, but it does not always lead to a higher randomness output

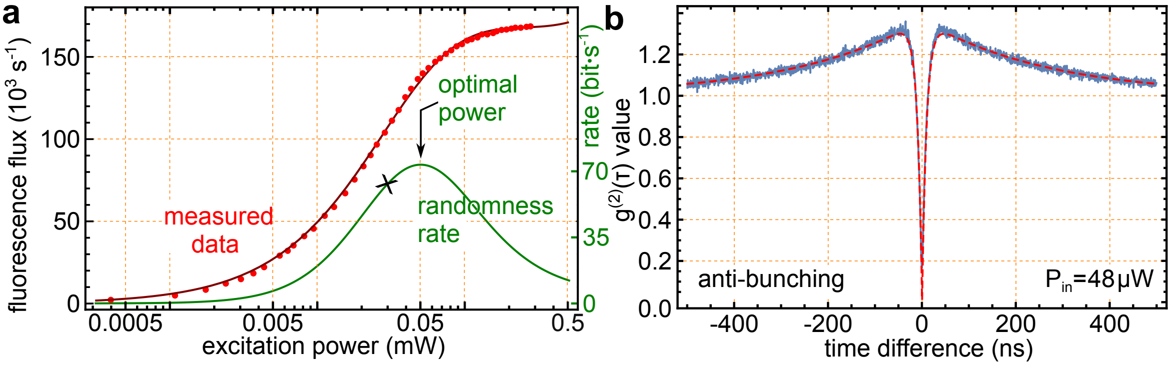


Figure 4.3: Randomness output speed with different laser powers. **a)** Saturation curve of the utilized NV center. Note that the non-trivial behavior at higher laser powers indicates that the NV center can not be considered a simple three-level system; it has more complex energy levels. The maximum output speed of the randomness generation (from the third model) is green. This curve forms because the antibunching curve gets narrower with increasing laser power. This implies that although more raw bits are generated per second, the overall area below the curve is reduced. The cross × in the green curve is the input laser power of our experimental data. **b)** The antibunching curve at the maximum output randomness output speed of the randomness generator. The bottom at τ=0 amounts to $g^{(2)}(0) = 0.15$. The figures are adapted from our publication [157].

speed. The reason is that the green area will reduce due to the antibunching curve getting narrower with higher laser power. The relationship between the randomness output speed and laser power is shown in Fig. 4.3.

The above equation can determine the fraction of extractable quantumness per raw bit.

Note that the quantum random bits generated in this area are supposed to be well-balanced, and the fraction of true randomness per raw bit is $r_{\text{rand}}/r_{\text{total}}$. This fraction is affected by the shape of the antibunching curve and $g^{(2)}_{\text{fit}}(0)$. An extreme case is when there are no single-photon events, such as when $g^{(2)}_{\text{fit}}(0)$ is unity. In this case, no quantum randomness is generated.

Since the fraction of true randomness per raw bit is $r_{\text{rand}}/r_{\text{total}}$, the remaining $1-r_{\text{rand}}/r_{\text{total}}$ bits are considered as classical noise, which is known by Eve. Correspondingly, $p_{\text{c}} = 1 - r_{\text{rand}}/r_{\text{total}}$, is the fraction of classical noise in per raw random bit. The conditional min-entropy can then be written as:

$$H_\infty(\mathbf{X}|\mathbf{Y}) = -\log_2\Big(p_{\text{c}} + (1-p_{\text{c}})\big(\sum_{\mathbf{y}} p(\mathbf{y}) \max_{\mathbf{x}}\{p(\mathbf{x}|\mathbf{y})\}\big)\Big) \,. \tag{4.8}$$

4.3 Experiment results analysis

In 7-day experimental run, we acquired 832 GiB data, corresponding to 55,796,707,904 raw bits. Then the number of zeros is 21,753,096,536 bits, and the number of ones is 34,043,611,368 bits. The integrated imbalance of the beamsplitter ratio amounts to probability $p(1)$=0.6101, $p(0)$=0.3899, which are indicated in Fig. 4.1e. This bias in the beamsplitter will largely decrease the usability of the generated random bits. We must post-process the raw randomness bits to get a uniformly distributed random sequence. As mentioned in Chapter 3, conditional min-entropy and randomness extractors are needed for this process.

In our first model, the conditional min-entropy $H_\infty(X|Y)$ is calculated by Eqn. (4.5), which gives us a conservative bound of the true randomness per raw bit. When considering the 11.5σ error bound, $H_\infty(X|Y)$ is 0.5559 bits, which means 0.5559 bits secured random number can be extracted per raw bit. With Toeplitz hashing extractor mentioned in Chapter 3, a uniformly distributed random sequence with 3.10×10^{10} bits can be extracted. Taking the refocusing periods into account, the output speed of pure quantum random bits is 5.10×10^4 bits per second.

By limiting the generated raw random data to single-photon events in the second model, we can guarantee independence of the random data from an uncorrelated background. Using Eqn. (4.6), with 11.5σ error bound, the extractable quantum randomness per raw bit amounts to 0.5168 bits. The total quantum random bits are 2.88×10^{10} bits for the whole raw bits. The random number output speed amounts to 4.74×10^4 bits per second when including the refocusing periods.

The difference between these two models indicates that some classical noise backgrounds might have been considered random events in the first model.

Next, we calculate the amount of quantum random bits from the third model.

The excitation power affects the fluorescence counts and the shape of antibunching curves. Subsequently, R_{rand} depends on the excitation power of the single quantum emitter. As shown in Fig. 4.3a, the green curve is the quantum randomness output speed; it depends on different excitation powers. The curve has an optimal excitation power. This is because, with increased excitation power, the count rate of different detectors increases while the shape of the antibunching curves becomes narrower. Thus the green part in Fig. 4.2 would become smaller. At the given excitation intensity of 26μW, the green part covers a time range of t =21.1 ns. When the excitation power changes, the start-stop event count rate will increase and decrease later. Thus, the quantum random bits output speed has an optimal operating point. This rate matches the single-photon emitter's saturation point for a simple three-level system.

Following Eqn. (4.8), with a very strict 11.5σ error bound, we compute the certifiable quantum randomness per raw bit as 3.746×10^{-4}. With this value for the Toeplitz hashing extractor, about 2.09×10^7 bits uniformly distributed random sequence can be extracted, and the random number generation speed is about 34.37 bits per second.

The NIST Statistical Test Suite results of all three models are shown in Fig. 4.4. The figure shows that the randomness quantified by different models all pass the test with a given significance level $\alpha = 0.01$. But their security is different from the above description. Although NIST test results cannot determine the privacy of the random sequence, they still provide a good measure of the extracted random sequence's mathematical structure. If the randomness sequence from the randomness extractor passes the NIST test, then the sequence is ready to use.

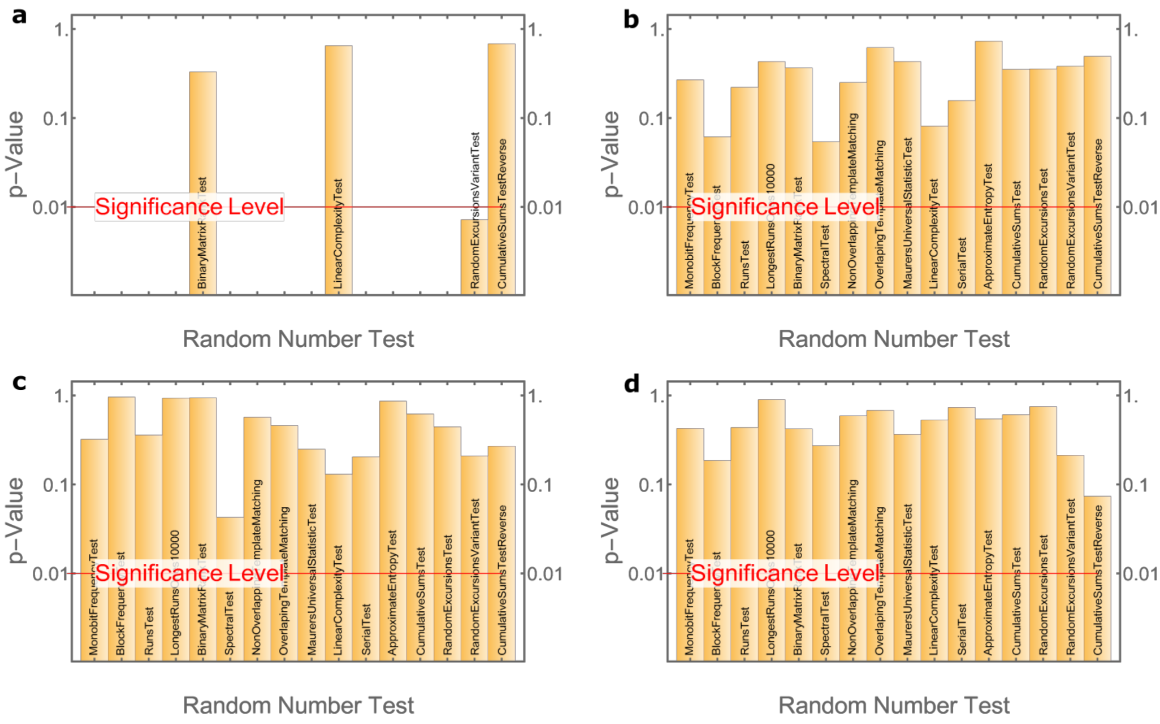


Figure 4.4: The NIST test results for raw bits and extracted bits with three models mentioned above. In all sub-figures, p values less than 0.001 are omitted. Figure **a** shows the NIST test results for one raw bits file (out of 179 files). The total bit length is about 298 Gbits. The raw bits clearly cannot pass the NIST test, even if it contains quantum randomness. Figure **b** to **d** shows the NIST test results for the extracted random bits based on three different models mentioned in this chapter. Although they all pass the test, they have different security levels, which depend on how the randomness is quantified in the corresponding models.

4.4 Conclusion

In conclusion, we have theoretically described and experimentally demonstrated a random bit generator based on a single-photon source. The single-photon source is based on a single defect NV center in a diamond. The generator is operated continuously over one week, and all detector events are recorded as time tags to be conveniently post-processed.

In the first model, the detection of raw random bits, which are associated with the two output ports of the beamsplitter, resulted in a raw-bit stream, and an entropy analysis for the raw random bitstream was presented. However, this model has some subtleties since the single-photon detection process is prone to technical effects such as the beamsplitter ratio, electrical dead times, and jitter. In a further analysis, the second model estimates the amount of unwanted and potentially untrusted background events by the antibunching effect of single-photons. In the third method, only tuple detection events are considered raw bits. The limitation is further reduced to auto-correlation values below unity and excluded from the uncorrelated background to guarantee the quantum nature of the source.

In the second and third models, the quantum input state only certifies the “quantumness” of the utilized light source. While the “decision”, the experimental outcome is still based on the fair-sampling assumption of the beamsplitter. This fair-sampling assumption can be lifted in multiple ways, shown in the subsequent chapters.

5 Randomness from a delayed-choice experiment

The antibunching effect of single-photons can be used to guarantee the authenticity of the light source and detect misalignment. However, the randomness still comes from the so-called fair-sampling assumption of the beamsplitter. In Chapter 2, we introduced the delayed choice experiment to eliminate the fair-sampling assumption. In this chapter, we illustrate our QRNG model to show how quantum randomness can be certified by wave-particle duality from the raw experimental data in the delayed-choice experiment performed by Jacques *et al.* [1].

5.1 Randomness certification protocol

For simplicity, the equivalent experiment scheme of the delayed-choice experiment in [1] is shown in Fig. 5.1.

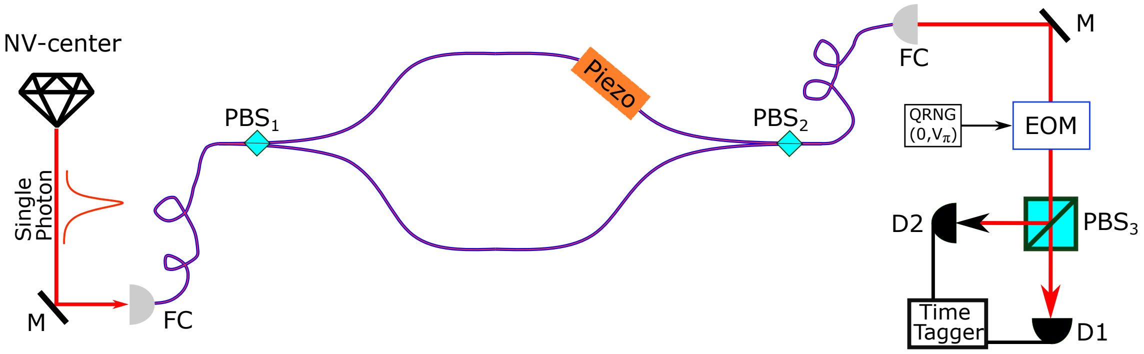


Figure 5.1: Simplified experimental scheme for the delayed choice experiment in [1]. The **Piezo** in the upper path is a discrete piezoelectric stack that can introduce a phase change for the two paths. **FC**: fiber coupler; **M**: mirror. The **Piezo** and PBS$_2$ are acting as the BS′ in Fig. 2.17.

In order to quantify the entropy of quantum randomness in the raw data from this delayed-choice QRNG, the "quantumness" in the data must be quantified first. As mentioned in

Chapter 2, the quantumness in the raw bits can be quantified by the interference visibilities of the interference patterns in the two detectors. This is because the interference only happens when the photon is in a superposition state of the two paths after it comes out from PBS$_1$.

The entropy of the randomness per raw bit is defined again by the conditional min-entropy $H_\infty(X|E)$, where X means the value of the raw bits, and E represents all the knowledge of the device known by the eavesdropper. $H_\infty(X|E)$ is defined as

$$H_\infty(X|E) = -\log_2\ p_{\text{guess}} \tag{5.1}$$

where p_{guess} is the maximum guessing probability of the events. Next, we need to establish the connection between p_{guess} and the interference visibility.

There are two detectors, D$_1$ and D$_2$, in our delayed-choice QRNG. The interference visibility of the two detectors is the same if the count rates in the two detectors are equal to each other. We consider a more general situation where the count rates of the two detectors are not equal. This situation is shown in Fig. 5.2. In this figure, the blue and green lines represent the count rates of the two detectors when no voltage is applied to the EOM. They are n_1 and n_2, and we assume $n_1 \geq n_2$. When the voltage is applied to the EOM, the interference patterns in each detector will be shown as dashed lines. From optical interference theory [162], we know that

$$\begin{aligned} &C_{\text{diff}} = n_1^{\max} - n_1^{\min} = n_2^{\max} - n_2^{\min}, \\ &n_1 = (n_1^{\max} + n_1^{\min})/2, \\ &n_2 = (n_2^{\max} + n_2^{\min})/2 \end{aligned} \tag{5.2}$$

According to Eqn. 2.7, the interference visibility for D$_1$ is $v_1 = \frac{C_{\text{diff}}}{2n_1}$, and for D$_2$ is $v_2 = \frac{C_{\text{diff}}}{2n_2}$. Both v_1 and v_2 are important in our protocol. Next, we show how to get the guessing probability of the raw bits with the given interference visibility v_1 and v_2.

Similar to previous chapters, we use the worst-case scenario to estimate the guessing probability. Suppose the density matrix of the state from the source is ρ,

$$\rho = z\,|\psi\rangle\,\langle\psi| + \frac{(1-z)\mathbf{I}_2}{2}, \tag{5.3}$$

where $|\psi\rangle = \alpha\,|H\rangle + \beta\,|V\rangle$, and $\mathbf{I}_2$, which represents a mix state, is a unity matrix in a 2-dimensional Hilbert space. When the photon passes the PBS$_1$, it has probability $\alpha\alpha^*$ to be transmitted and $\beta\beta^*$ to be reflected. The mixed state part $\mathbf{I}_2$ represents all the untrustworthy

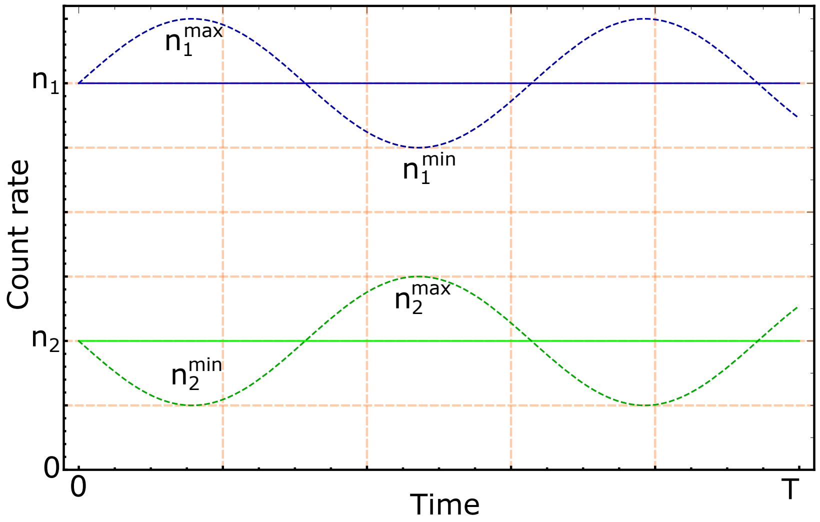


Figure 5.2: Different count rates of the two detectors in a MZI. When the photons from two paths are interfering, the count rates in the two detectors changes and the interference patterns are shown. Although the count rates of the two detectors are not the same, they can still interfere, and the relative count rates differences of the two detectors are the same, which means $n_1^{\max} - n_1^{\min} = n_2^{\max} - n_2^{\min}$.

parts of the detector events, including but not limited to the dark counts, the external sources, and the misalignment of the devices. The guessing probability of $|\psi\rangle$ in which path is $\alpha\alpha^*$ (assuming $|\alpha| \geq |\beta|$), and of mixed state $\mathbf{I}_2$ is 1. Combining the two parts, the total guessing probability is

$$p_{\text{guess}} = z\alpha\alpha^* + (1-z) \tag{5.4}$$

Whether the count rates in the two detectors are equal or not, each detector can get the incoming photons from two possible sources. One source is from the collapse of the pure state $|\psi\rangle$, and the other source is the mixed state $\mathbf{I}_2$. Among all the possible source combinations, the following situation can maximize the guessing probability with given interference visibilities v_1 and v_2: All photon events in D_1 are the transmitted photons from state $|\psi\rangle$, and all the $\mathbf{I}_2$ and reflected photons of $|\psi\rangle$ only click at D_2. Then, we have the following equations

$$\begin{aligned} z\alpha\alpha^* &= \frac{n_1}{n_1+n_2} \\ z\beta\beta^* + (1-z) &= \frac{n_2}{n_1+n_2} \end{aligned} \tag{5.5}$$

Considering $\alpha\alpha^* + \beta\beta^* = 1$, $v_1 = \frac{C_{\text{diff}}}{2n_1}$, and $v_2 = \frac{C_{\text{diff}}}{2n_2}$, we get the solution

$$\begin{aligned} \alpha &= \frac{1}{\sqrt{1+v_1}}, \\ \beta &= \frac{v_1}{\sqrt{1+v_1}}, \\ z &= \frac{v_2(1+v_1^2)}{v_1+v_2} \end{aligned} \tag{5.6}$$

Then with Eqn. 5.4, the guessing probability is derived as

$$p_{\text{guess}} = 1 - \frac{v_1 v_2}{v_1 + v_2} v_1 \tag{5.7}$$

The lower bound of this guessing probability is 0.5. This bound can only be reached when the photons after the PBS$_1$ can be described by a pure quantum state $(|H\rangle + |V\rangle)/\sqrt{2}$, which results in $v_1 = v_2 = 1$. In this case, with the absence of voltage on the EOM, photons will hit the two detectors randomly, and random numbers with guessing probability 0.5 are generated. When $0 < v_{1,2} < 1$, the photons are not in the pure state $(|H\rangle + |V\rangle)/\sqrt{2}$, and quantumness is partially involved in the measurement. In one extreme case, the visibility v_1 and v_2 can be 0. This means that after the photons pass the PBS$_1$, they are not in superposition states, so the interference is not happening. And the guessing probability is 1 in this situation, which means no quantum randomness is generated since the click events in the two detectors are potentially controlled by the eavesdroppers.

According to Eqn. 5.1 and Eqn. 5.7, the relationship between the two interference visibilities and the min-entropy per raw bit is shown in Fig. 5.3.

If there is a sampler behind PBS$_1$, the photons will hit the sampler, the superposition state will collapse, and there will be no interference pattern in D$_1$ and D$_2$ when voltage is applied to the EOM. In other words, if the interference pattern is detected, it means the photon arrives at detectors D$_1$ and D$_2$ as a superposition state, and it interferes with itself. In the delayed-choice QRNG, the decision of applying or no-applying voltage to the EOM is unknown by PBS$_1$, which means if the interference pattern is observed in this QRNG, the randomness generation is free from fair-sampling assumption.

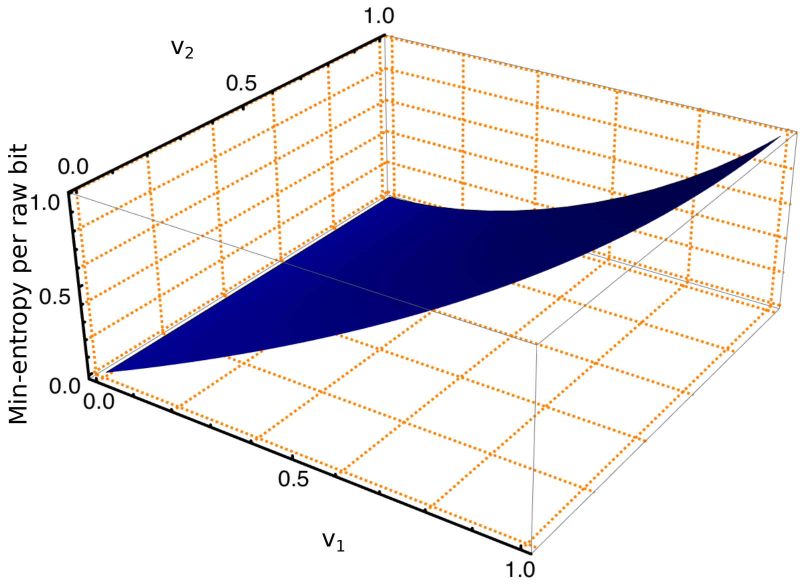


Figure 5.3: The relationship between $v_{1,2}$ and the min-entropy per raw bit. We assume $v_1 \leq v_2$ here. As v_1 and v_2 increase, the entropy per raw bit also rises.

5.2 Confidence level analysis

As a standard process, the relationship between the confidence level of the model and the interference visibility should be analyzed due to the finite size of experimental data. From the Eqn. 5.1 and Eqn. 5.7, we have

$$H_\infty(X|E) = f(v_t). \tag{5.8}$$

where

$$\begin{aligned} f(v_1, v_2) &= -\log_2 p_{\text{guess}} \\ &= -\log_2(1 - \frac{v_1 v_2}{v_1 + v_2} v_1) \\ &= -\log_2(1 - v_t) \end{aligned} \tag{5.9}$$

where $v_t = \frac{v_1 v_2}{v_1 + v_2} v_1$.

Define the confidence level as $1 - \delta$, and the error bound of v_t as ϵ_t. Then the min-entropy per raw bit with confidence level $1 - \delta$ is

$$H_\infty(X|E) = f(v_t - \epsilon_t). \tag{5.10}$$

When considering all the events n, the total min-entropy of the experimental data is $H_\infty(R|E) = nH_\infty(X|E) = nf(v_t - \epsilon_t)$. Next, we derive the relationship between $1-\delta$ and ϵ_t. Similar to previous chapters, their relationship can be derived from Hoeffding's inequality [163]. For single event, v_t can be defined as v_{t_i}

$$v_{t_i} = \frac{v_{1_i} v_{2_i}}{v_{1_i} + v_{2_i}} v_{1_i} \tag{5.11}$$

Since the interference visibility of each event is only decided by the superposition status of the incoming photon, v_{t_i} can be treated as independent of each other. Since $0 \le v_1, v_2 \le 1$, so $0 \le v_{t_i} \le 0.5$. Then according to Hoeffding's inequality [163], we have

$$\delta = \exp\left(-8n\epsilon_t^2\right) \ge Pr\left[|v_t - v_{t_r}| \ge \epsilon_t\right] \tag{5.12}$$

where v_{t_r} means the expected value from the experimental setup. This equation means that the v_{t_r} can be lower than v_t up to error ϵ_t with small probability δ. Then the entropy of randomness can be quantified with confidence level $1-\delta$.

5.3 Result analysis

In this section, we apply our delayed-choice QRNG model to the delayed-choice experiment performed by Jacques *et al.* [1] to show how randomness can be bounded in this experiment.

The interference visibility of detector D$_1$ and D$_2$ in the experiment is reported as $v_1 = v_2 = 0.94$. Then we have $v_t = 0.4418$. This v_t value is obtained from $n = 2,600$ photons. With 99% confidence level, we have $v_{t_r} = 0.4269$. According to Eqn. 5.9, the min-entropy of per raw bit is $H_\infty(X|E) = 0.8031$. When no voltage V$_\pi$ is applied to the EOM, the setup generates quantum randomness. The total photon count rate is about 1,400 counts per second, so the randomness output speed is around 1,124 bits per second.

5.4 Conclusion

In this chapter, we connect the counter-intuitive phenomenon of a delayed-choice experiment with quantum randomness generation. We demonstrate how to build a QRNG from the delayed-choice experiment, and a QRNG protocol based on interference visibility is introduced to certify quantum randomness from the raw bits.

Compared with the single-photon QRNG mentioned in Chapter 4, the entropy of the randomness can not only be quantified without classical noise but also without fair-sampling assumption in the beamsplitter. We then apply our model to the delayed-choice experiment in [1]. The interference visibility reported in this paper is 0.94, and with a 99% confidence level, the min-entropy per raw bit is 0.8031, and the random number output speed is 1124 bits per second.

6 Randomness from a loophole-free Bell test

In this chapter, we show how to build a quantum random number generator from a loophole-free Bell test [62], in both DI and SDI approaches. For the DI approach, we use the earlier analysis from [82] to quantify the entropy in the data. We introduce a RSP dimension witness for the SDI approach. This allows us to get a higher bound of min-entropy per event. The two schemes are compared in Fig 6.1.

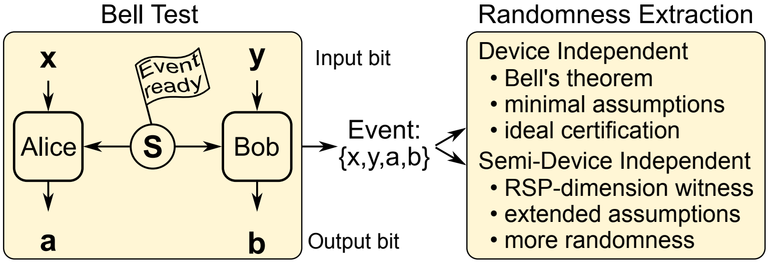


Figure 6.1: Two different ways to bound the quantum randomness in a Bell test. A Bell test involves two physically separated experimental systems, with two given input bits x, y, to generate two binary outcomes a, b. The Bell correlation value S allows a DI scenario to bound the min-entropy of the randomness; another SDI scenario is to extract randomness when RSP-dimension witness is utilized. The figure is adapted from [3]

The QRNG based on the Bell test is a collaboration work between the group of Prof. Harald Weinfurter (Ludwig Maximilian University of Munich) and Prof. Jörg Wrachtrup (University of Stuttgart). The Ludwig Maximilian University of Munich performed the Bell test, and their work is summarized in [62]. The University of Stuttgart analyzed the experimental data and built models to quantify the entropy of randomness in the raw data, and this work is summarized in [3].

6.1 Experimental setup

The bipartite loophole-free Bell test performed here [62] is to test Bell's theorem in the form of the CHSH inequality [73]. The detailed description of their experimental setup can be found at [62], but for the sake of consistency and clarity within this chapter, we rephrase it in this section.

In this experiment, Alice and Bob each operate an atom trap for a single rubidium atom. The two traps, separated by 398 m, are independently operated, comprising their laser system and control electronics. The atomic qubits are encoded in the $m_F = \pm 1$ Zeeman sub-level of the $5S_{\frac{1}{2}}, F = 1$ ground state, with $\left|\uparrow\right\rangle_z$ corresponding to $m_F = +1$, $\left|\downarrow\right\rangle_z$ corresponding to $m_F = -1$.

To create the entangled atom-photon pairs, each atom is excited to the $5P_{\frac{3}{2}}, F' = 0, m_F = 0$ state via a short laser pulse. The subsequent spontaneous emission yields a photon whose polarization is entangled with the atomic qubit state. Both photons are then coupled into single-mode fibers and guided to a Bell state measurement (BSM) setup. Two-photon interference on a fiber BS and photon polarization analysis is employed to project the photons on two of the four Bell states. The photonic measurement heralds the creation of one of the entangled atom state $|\Psi^{\pm}\rangle = 1/\sqrt{2}\,(\left|\uparrow\right\rangle_x \left|\downarrow\right\rangle_x \pm \left|\downarrow\right\rangle_x \left|\uparrow\right\rangle_x)$, where $\left|\uparrow\right\rangle_x = (1/\sqrt{2})\,(\left|\uparrow\right\rangle_z + \left|\downarrow\right\rangle_z)$ and $\left|\downarrow\right\rangle_x = (i/\sqrt{2})\,(\left|\uparrow\right\rangle_z - \left|\downarrow\right\rangle_z)$.

After an entangled pair is created between the two sides (Alice and Bob), they start a fast atomic state measurement process based on state-selective ionization and subsequent detection of the ionization fragments. The measurement setting is determined by the polarization of a laser pulse exciting the atom before ionization. Each party employs a QRNG outputting freshly generated random bits on demand to decide the choice of the setting. The total time needed from the generation of the input pair x or y to receiving the output pair a or b is less than 1.1 μs, together with a separation of the atom traps of 398 m. This enables space-like separation of the measurements [81]. In total, 55,568 rounds were recorded, 27,885 with the $|\Psi^+\rangle$ prepared and 27683 with the $|\Psi^-\rangle$. The experimental setup is shown in Fig. 6.2

The Bell state measurement is performed on two photons from Alice's and Bob's sides, and high visibility of the two-photon interference of the photons is necessary. This is ensured by precise adjustment and synchronization of the excitation laser pulse and other experimental parameters. The combined photon collection and detection efficiency for

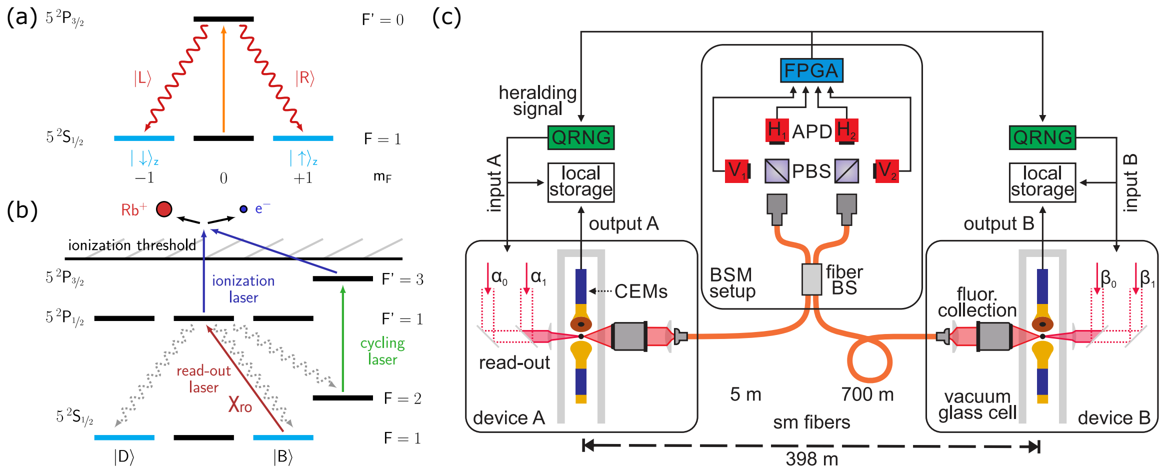


Figure 6.2: Experimental scheme of our loophole-free Bell test. (a) Simplified scheme of atom-photon entanglement generation: rubidium atoms in the $5S_{\frac{1}{2}}, F = 1, m_F = 0$ state are optically excited to $5P_{\frac{3}{2}}, F' = 0, m_F = 0$. The subsequent spontaneous decay results in the entangled atom-photon state $|\Psi_{AP}\rangle = \frac{1}{\sqrt{2}}\left(|L\rangle\,|\downarrow\rangle_z + |R\rangle\,|\uparrow\rangle_z\right)$. (b) State selective ionization scheme: depending on the polarization χ_{ro} of the read-out laser pulse a selected superposition of the $m_F = \pm 1$, Zeeman ground states is excited to the $5P_{\frac{1}{2}}, F' = 1$ a level that is ionized by a second laser pulse. During this process, the excited state can decay to the $F = 1$ and $F = 2$ ground levels before ionization (gray arrows). The population in $F = 2$ is excited by a third laser pulse to $5P_{\frac{3}{2}}, F' = 3$ from which it is ionized, while the decay to $F = 1$ reduces the fidelity of the measurement process. (c) Sketch of the experimental setup: The two devices, A and B, independent apparatuses for trapping single atoms, are separated by 398 m. Entanglement between the two single atoms is created by first entangling each atom with a photon that is then coupled into single-mode fiber and guided to a photonic BSM setup. There, the two photons are overlapped on a fiber beam splitter, and subsequent polarization analysis of the photons projects the atoms in an entangled state. The two-photon detection is analyzed with a field programmable gate array (FPGA) that sends a heralding signal to each device in case of a successful entanglement generation. This signal triggers quantum random number generators to generate input bits, determining polarization of the read-out laser pulse χ_{ro} in the atomic state measurement. The measurement results, registered by the channel electron multipliers, are recorded together with the input bits in local storage devices. The figure is adapted from [62].

device A is $\eta_A \approx 1.6 \cdot 10^{-3}$ and for device B is $\eta_B \approx 0.8 \cdot 10^{-3}$. These two efficiencies lead to a total success probability $6.4 \cdot 10^{-7}$ to create atom-atom entanglement. With an effective excitation repetition rate of ≈ 50 kHz, an average rate of 1 to 2 of entangled atom-atom pairs per minute is achieved during the experiment [62].

In case of the successful creation of entanglement, a signal is sent to each device, triggering the measurements on the atomic qubits. These measurements are based on fast state-selective ionization and subsequent detection of the ionization fragments. A read-out laser pulse excites, depending on its polarization χ_{ro}, a certain superposition of the qubit states $|B\rangle$ to the $5P_{\frac{1}{2}}$ level, which is ionized with a second laser pulse (Fig. (6.2) (b)). The ionization fragments are detected with two-channel electron multipliers[164], yielding the results +1 (detection of at least one fragment) or -1 (no detection). This fast and highly efficient atomic state measurement yields a result within 1.1μs, including the choice of the measurement setting with fidelity of 0.97 [62]. To ensure the independence of the measurement settings, the read-out polarizations are chosen locally by quantum random number generators. This, in combination with the fast atomic state measurement and a distance of 398 m between the devices, allows for space-like separation of the measurements and thus ensures independence of the local measurements.

6.2 DI randomness

The quantum randomness generation can be connected to the violation of Bell inequalities [165, 82]. The violation of Bell inequalities guarantees that the measurement results are not pre-determined and must be from the entangled system, which possesses intrinsic randomness. For the loophole-free best test, the Bell inequalities are violated in a DI manner, which means the experimental device is not trusted, and the violation is only calculated by the measurement results. This implies that the randomness generated from loophole-free bell tests is also DI randomness.

Although the randomness certified by Bell's theorem can be device independent, some additional assumptions are still needed to bound the randomness in this model [82]: (1) the remote parties perform local, and independent measurements on their ideally space-like separated (=perfectly shielded) devices; (2) the measurement settings (x,y) are not determined beforehand and are unpredictably chosen; (3) the measurement process is described by quantum mechanics. In a loophole-free Bell test, assumption (1), which is required for a loophole-free Bell test, is fulfilled. Assumption (2) means that the $i-th$ input x_i and y_i are revealed to the experimental devices until the $i-th$ run of the experiment.

In [82], the marginal guessing probability $\max_{ax} p(a \mid x)$ had been linked to the correlation value S of the CHSH inequality.

$$\max_{ax} p(a \mid x) \leq \frac{1}{2}\left(1+\sqrt{2-\frac{S^2}{4}}\right) \tag{6.1}$$

This equation allows us to bound the entropy of the output data to 1 bit per event when $S = 2\sqrt{2}$ by min-entropy $H_\infty = -\log_2 \max_{ax} p(a \mid x)$. From Eqn. 6.1, we know that for the certification of the true randomness from Bell test data, Bell inequalities must be violated. Unfortunately, the loophole-free Bell experiments [61, 80, 79, 62] did not reach the maximally allowed values for quantum mechanics. This means only a small amount of randomness per event can be bounded in the DI manner. Sometimes no randomness can be certified by Bell's theorem [1]. However, this situation changes when additional assumptions are introduced to leave the DI scenario. In order to build the experimental devices, it is necessary to have some knowledge about how they function, e.g., the devices are error-prone but not maliciously built. This knowledge about the experimental devices allows for a higher min-entropy bound of the quantum randomness per event for the same experimental data.

6.3 SDI randomness

In order to build an SDI protocol, a *dimension witness* can be used. The concept of dimension witness was first introduced in [166]. After this paper, many studies have been performed on this concept [139, 167, 168, 169, 83].

6.3.1 2-Dimensional quantum representation proof for our system

Before applying the dimension witness to the Bell test scenario, especially the CHSH scenario of our experiment [62], we first show that the experiment admits a 2-dimensional quantum representation [166, 139, 83].

In our loophole-free Bell test, there are two binary inputs x, y and two binary outputs a, b. When Alice perform her measurement, the entangled state on Bob's side will randomly collapse into a specific state. Since Alice has two different measurement settings, and each one has two different measurement results, Bob's side will get four quantum states when she does her two measurements randomly multiple times. These four quantum states on Bob's side can be represented as x', and $p(b|x', y) = p(ab|xy)/p(a|x, y)$.

Notice that, $p(ab|xy) = p(a|xy)p(b|x, a, y)$, therefore $p(b|x', y) = p(b|x, a, y)$. So we can treat Alice's outputs as the input parameter [169]. This means that Alice's input x and result a together can be treated as the state labels x' for Bob. Next, we need to prove that $p(b|x', y)$

[1] This may be caused by (i) the finite amount of experimental events or (ii) Bell inequality is not violated

can be re-written as $p(b|x',y) = \text{Tr}\left(\rho_{a|x} M^B_{b|y}\right)$, where $\rho_{a|x}$ is the state prepared on Bob's side when Alice performs her measurement x and gets a result a. $M^B_{b|y}$ is the measurement operator on Bob's side.

In the quantum formalism, the state shared by the two parties ρ_{AB} is defined on $\mathcal{H}_A \otimes \mathcal{H}_B$, and the measurement performed by Alice $M^A_{a|x}$ acts on $\mathcal{H}_A$ and the one performed by Bob $M^B_{b|y}$ acts on $\mathcal{H}_B$. The binary input and output in the CHSH scenario allow us to describe the output probabilities $p(ab|xy)$ with 2-dimensional $\mathcal{H}_A$ and $\mathcal{H}_B$ (isomorphic to $\mathbb{C}^2$, a 2-dimensional complex coordinate space). With this, the probability for a measurement outcome b at Bob's side depending on the input y as well as Alice's measurement x and result a is

$$\begin{aligned} p(b|x',y) &= \frac{p(ab|xy)}{p(a|x,y)} \\ &= \frac{\text{Tr}\left[\rho_{AB}\left(M^A_{a|x} \otimes M^B_{b|y}\right)\right]}{\text{Tr}\left[\rho_{AB} M^A_{a|x} \otimes \mathbf{1}\right]}, \end{aligned} \tag{6.2}$$

where $\mathbf{1}$ is an identity operator. This can be interpreted as Alice preparing a state at Bob's side with her measurement:

$$\begin{aligned} &\text{Tr}\left[\rho_{AB}\left(M^A_{a|x} \otimes M^B_{b|y}\right)\right] \\ =&\text{Tr}\left[\underbrace{\left(\rho_{AB}(M^A_{a|x} \otimes \mathbf{1})\right)}_{Alice} \otimes \underbrace{\left(\rho_{a|x} M^B_{b|y}\right)}_{Bob}\right]. \end{aligned} \tag{6.3}$$

The left part represents the measurement on Alice's side, and in the right part, $\rho_{a|x}$ is the state prepared on Bob's side after Alice's measurement. Inserting this in Eqn (6.2) directly leads to

$$p(b|x',y) = \text{Tr}\left[\rho_{a|x} M^B_{b|y}\right]. \tag{6.4}$$

Since $\rho_{a|x}$ and $M^B_{b|y}$ are acting on $\mathbb{C}^2$. Subsequently, $p(b|x',y)$ admits a 2-dimensional quantum representation. This shows that we can use the two-dimensional dimension witness [83] to quantify the quantumness in our experiment.

6.3.2 Applying dimension witness to CHSH scenario

Different dimension witnesses can be used in a 2-dimensional quantum representation. The dimension witness we used here is introduced in Chapter 2. We chose this dimension witness because this nonlinear dimension witness can be used for non-convex sets and is robust to technical imperfections. Most importantly, it can be used to certify quantum randomness [83]. It is defined as

$$W = \begin{vmatrix} p(1|0,0) - p(1|1,0) & p(1|2,0) - p(1|3,0) \\ p(1|0,1) - p(1|1,1) & p(1|2,1) - p(1|3,1) \end{vmatrix}, \tag{6.5}$$

where $p(b|x',y)$ is defined in Eqn. (6.4), and the result b is chosen as "1" in the above definition. The definition of the dimension witness here is the same as in [83], but the state x' differs. Here, the state x' is on Bob's side, but the projective measurement performed by Alice on the entangled state completes its preparation, so the state x' is remotely prepared [170, 171]. In order to emphasize this difference, we name it as RSP-dimension witness.

RSP is a special case of quantum teleportation [172, 170]. In such a scheme, two parties, Alice and Bob, share an entangled state (ρ_{AB}). If Alice now wants to prepare a certain state at Bob's side, Alice measures her part of the entangled state. Depending on the measurement outcome, Bob's part collapses into a certain state. Since a quantum process generates Alice's measurement result, Bob's state will randomly collapse to a certain state. If a deterministic state preparation is desired, applying a unitary transformation, which depends on Alice's outcome, to Bob's state is necessary. However, it is not necessary for our protocol.

Briefly speaking, in RSP-dimension witness, Alice performs, depending on the input x, one of two projective measurements, $M^A_{a|x}$ on the entangled state ρ_{AB}, and one of four different states $\rho_{a|x}$ is prepared on Bob's side. Then with Bob's measurements y and measurement results b, the RSP-dimension witness W_B for Bob's side can be constructed. Similarly, W_A can be constructed for Alice's side.

In a Bell test, the two parties, Alice and Bob, share an entangled state. This entangled state is symmetric between Alice and Bob in general. This means that changing Alice and Bob's roles will not affect the analysis results. This indicates that $W_{\text{rsp}} = W_A = W_B$, but Alice's and Bob's experimental devices are not exactly the same. They are affected by different classical noises, which result in the difference between W_A and W_B. To get a more conservative bound of quantumness in the measurement result, we define $W_{\text{rsp}} = \min\{W_A, W_B\}$. Then we use

W_{rsp} as the RSP-dimension witness in the following model. The RSP-dimension witness captures the quantumness of the state preparation and measurements in our Bell test. If the preparations are classical, one has $W_{\text{rsp}} = 0$, while a quantum preparation and measurement leads to $0 < W_{\text{rsp}} \leq 1$.

Although S and W_{rsp} are based on the same experimental data, they are not directly affected by each other: S cannot be used to calculate the value of W_{rsp}, it only affects the lower bound of the W_{rsp}. For example, when $S = 2$, $W_{\text{rsp}} \in [0, 1]$, and when $S = 2\sqrt{2}$, $W_{\text{rsp}} = 1$.

Before using the RSP-dimension witness to bound the min-entropy of the randomness generated in the experiment, we discuss the required assumptions. The above (DI-) assumptions (1, 2, 3) still hold. Besides, there are some extra assumptions [4]: (4) the information in the measurement results of each side is contained in a two-dimensional quantum subspace; (5) the system is memoryless, and subsequent outcomes are not directly correlated.

In the RSP-dimension witness model, the state preparation and measurement must be independent of each other. This means for W_B, the states x' on Bob's side are remotely prepared by Alice's quantum measurements on the entangled states, which cannot be affected by Bob's device or measurement, so the states x' are independent of Bob's device and measurement setting y. This independence requirement is naturally fulfilled by our loophole-free Bell test [62]. The prepared states x' might be affected by Alice's experimental device, but that is not a concern in our model. Because if x' is affected by the device on Alice's side, it will not be properly prepared on Bob's side, and the value of the RSP-dimension witness will be decreased. Therefore, the independence of x' is quantified by the value of the RSP-dimension witness. Assumption (1) also implies that the experimental devices do not have any pre-established correlations among each other; this also indicates that the devices used to generate the input strings x, y are not correlated with the measurement devices. Subsequently, x, y can be pseudo-random numbers as long as they are independent of each other and the measurement apparatus.

Assumption (4) means that the information contained in the measurement result when measuring x' does not exceed 1 bit. A possible violation would be that the information about x' is duplicated by or correlated with extra qubits. An entangled state shared between Alice and Bob has two different measurement results on each side with one measurement setting–a qubit can describe it. This does not mean the entangled state shared between Alice and Bob has to be confined in a two-dimensional Hilbert space. It only means that with the measurement performed by Alice or Bob on it, the results information can be fully described by a qubit.

For our loophole-free Bell test, the state preparation in the RSP-dimension witness is independently accomplished by two sides: one side performs the measurement, and the other side gets the prepared state simultaneously. Under space-like separation, the measurement performed on one side is outside the light cone of state preparation on the other side. Thus, the state preparation devices cannot send extra qubits of the prepared states to the measurement devices without lowering the values of W_{rsp} [2]. As long as $W_{\text{rsp}} > 0$, the remote measurements exceed a classical correlation.

In the experiment, Alice and Bob perform measurements independently to their shared entangled state. Alice performs measurement according to x and Bob according to y, then Alice gets result a and Bob b. The measurement series x is prepared by Alice, and the result a is generated by a quantum process. Both x and a are independent of the experimental device, otherwise, when Alice applies the SDI model, the dimension witness W_{rsp} will decrease, and less randomness can be extracted.

With the assumption being settled, the guessing probability can now be derived. Since the two binary inputs, x and y, are independent of each other, and in the experiment, different choices of measurement settings are uniformly random sequences. Thus each combination of x and y occurs with probability $1/4$ [4]. Then, the guessing probability p_{guess} of $p(ab|xy)$ is (more intermediate steps are shown in Appendix)

$$\begin{aligned}
&p_{\text{guess}}(ab|xy) \\
&= \frac{1}{4}\sum_{x,y} \max_{a,b} p(ab|xy) \\
&\leq \max_{x,a} p(a|x) \frac{1}{2} \sum_{y} \max_{x,a,b} p(b|(x,a),y) \\
&\leq \left(\frac{1+\sqrt{1-W_{\text{rsp}}^2}}{2} \right) \frac{1}{2} \left(1 + \sqrt{\frac{1+\sqrt{1-W_{\text{rsp}}^2}}{2}} \right) .
\end{aligned} \tag{6.6}$$

[2] This means the extra qubits sent out by the state preparation devices can be treated as classical noise in this situation.

The equation of guessing probability $p_{\text{guess}}(ab|xy)$ from W_{rsp} is not the same as the one from [4]. The difference is caused by $\max_{x,a} p(a|x)$, which represents the quantum measurement from the state preparation process.

The conditional min-entropy $H_\infty(AB|XY)$ in this situation is $H_\infty(AB|XY) = -\log_2 p_{\text{guess}}(ab|xy)$. This equation allows us to bound the min-entropy of the randomness in the Bell test data in an SDI manner. The randomness per event from our RSP-dimension witness model is depicted in Fig. (6.3). Compared to [4], the introduction of quantum measurements in the state preparation process gives us a significant increase in bounding randomness in our experimental data. For instance, the maximum min-entropy of the quantum randomness in our model is 1.23 bits per event, which is much larger than the previous dimension witness model [4].

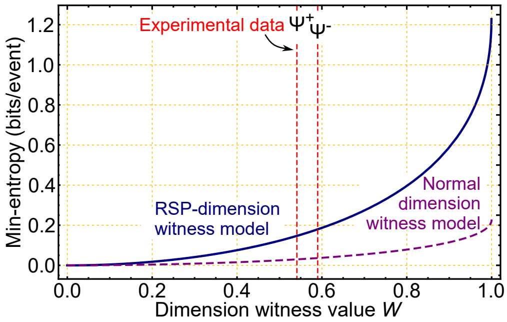


Figure 6.3: Output randomness utilizing the dimension witness. The nonzero RSP-dimension witness W_{rsp} gives us a new perspective to bound the min-entropy of randomness in the experimental data. In this figure, the blue curve displays the randomness bounded by the W_{rsp}, while the dashed purple curve represents the randomness bounded by the previously defined dimension witness [83]. Clearly, the combination of remote state preparation and the dimension witness increases the bound of randomness per event compared to a normal dimension witness QRNG model in [4]. The figure is adapted from [3]

The above RSP-dimension witness model can bound more quantum randomness in the Bell test from a different perspective, and only a few additional general assumptions are required for this. Moreover, when S is below the classicality bound 2, the W_{rsp} can still be larger than 0. See the example below.

The Bell inequality might not be violated in a practical Bell test because of the imperfect measurements or entangled states. In such a case, no randomness can be quantified by earlier models [82, 173, 174, 75, 74]. With the RSP-dimension witness model, randomness in the experimental data can be bounded without using Bell's theorem. For example, suppose Alice and Bob share one Bell state, then they measure it with two identical measurement settings $\widehat{x}$ and $\widehat{z}$ at each side (which corresponds to the BBM92 quantum key distribution scheme

[42]). Bell inequalities will not be violated in this case, but the min-entropy of quantum randomness is 1.23 bits per event data from W_{rsp}.

This example shows that by rotating Alice and Bob's measurement bases in the same plane of a Bloch-sphere, the RSP-dimension witness is not changed, and it also demonstrates the robustness of our SDI protocol. This shows us that quantum randomness in the Bell test data can be quantified without using Bell's theorem.

We further consider an example with a 2-qubit Werner state (6.7),

$$\rho_z = z|\Psi^+\rangle\langle\Psi^+| + \frac{1-z}{4}\mathbf{I}\,, \tag{6.7}$$

where $0 \leq z \leq 1$, is the noise parameter. The relationship between W_{rsp} and S can be derived for this state. On the one hand, the relationship between z and S is $S = 2\sqrt{2}z$. On the other hand, following [83], the relationship between z and the RSP-dimension witness is derived as $W_{\text{rsp}} = z^2$. Subsequently, the relationship between W_{rsp} and S is calculated as $W_{\text{rsp}} = S^2/8$. From this relationship, we can also see that W_{rsp} is nonzero when $0 < S \leq 2$. This shows again that randomness in the Bell test data can be certified without using Bell's theorem.

Our SDI model based on dimension witness and the DI model in [82] both utilize quantum correlations to bound the min-entropy of quantum randomness. The DI model uses the correlation between the measurement results and the measurements to form a CHSH inequality, the violation of CHSH inequality guarantees that the quantum randomness can be certified. In the SDI model, the correlation between Alice and Bob's measurement results is not quantified by Bell's theorem but by our RSP-dimension witness. This RSP-dimension witness can quantify the quantum correlation, which cannot be quantified by the CHSH inequality, such as the case in BBM92 scenario and 2-qubit Werner state.

6.4 Randomness extraction

In the randomness extraction process, the confidence level of the model and the error of hashing functions are introduced because of the finite data size. For the DI model, the confidence level analysis is in the supplementary of [82]. Here we only describe the confidence level details of our RSP-dimension witness model.

6.4.1 Confidence level of SDI randomness protocols

The SDI randomness in per event data is quantified by the RSP-dimension witness W_{rsp} $H_\infty(AB|XY) \geq f(W_{\text{rsp}})$. W_{rsp} has error ϵ_w, and the confidence level is $1-\delta$, this means

$$H_\infty(AB|XY) \geq f(W_{\text{rsp}} - \epsilon_w), \tag{6.8}$$

then all the SDI randomness $H_\infty(R|W_{\text{rsp}})$ in the n events experimental data can be calculated as

$$H_\infty(R|W_{\text{rsp}}) \geq nf(W_{\text{rsp}} - \epsilon_w). \tag{6.9}$$

Hoeffding's inequality can derive the relationship between ϵ_w and δ. In order to apply Hoeffding's inequality [163], we need to define W_{rsp_i} for a single event. Considering the definition of W_{rsp}

$$W_{\text{rsp}} = \begin{vmatrix} p(1|0,0) - p(1|1,0) & p(1|2,0) - p(1|3,0) \\ p(1|0,1) - p(1|1,1) & p(1|2,1) - p(1|3,1) \end{vmatrix}. \tag{6.10}$$

The relationship between W_{rsp} and W_{rsp_i} should be

$$W_{\text{rsp}} = \frac{1}{n}\sum_i W_{\text{rsp}_i}. \tag{6.11}$$

W_{rsp_i} could be defined as

$$W_{\text{rsp}_i} = \begin{vmatrix} p_i(1|0,0) - p_i(1|1,0) & p_i(1|2,0) - p_i(1|3,0) \\ p_i(1|0,1) - p_i(1|1,1) & p_i(1|2,1) - p_i(1|3,1), \end{vmatrix} \tag{6.12}$$

where $p_i(1|x,y)$ is the expected probability distribution for single event. Since the devices have no memory, the expected dimension witness W_{rsp_i} of each experimental run is independent of each other. For each W_{rsp_i}, we have $|W_{\text{rsp}_i}| \leq 1$. According to Hoeffding's inequality [163], the relationship between confidence level $1-\delta$ and the error bound ϵ_w is

$$\delta = \exp\left(\frac{-n\epsilon_w^2}{2}\right). \tag{6.13}$$

From this equation, the error bound ϵ_w can be calculated with a given confidence level.

$H_\infty(R|W_{\text{rsp}})$ in Eqn. 6.9 is the bound of quantum randomness in the experimental data. To get a uniformly distributed bit string from the experimental data, randomness extractors need to be used.

6.4.2 Randomness extractors

The extractor we choose is the Toeplitz hashing extractor. For the details of this extractor, see Chapter 3.

The raw random bits in our model are from two sides, Alice and Bob. The randomness from the two sides could exceed 1 bit per event data. Thus, extracting the bounded quantum randomness using a standard construction of extractors, which takes the form $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$, is not possible. We need to use two extractors to extract the randomness in the raw random bits with our model.

For the joint outcome experimental events, the SDI randomness per experimental run is

$$H_\infty(AB|XY) = -\log_2 \max p(ab|xy). \tag{6.14}$$

Note that

$$\max p(ab|xy) \leq \max p(a|x) \max p(b|x,a,y), \tag{6.15}$$

which means

$$H_\infty(AB|XY) \geq H_\infty(A|X) + H_\infty(B|XAY). \tag{6.16}$$

Therefore, we can bound the random strings from each side and combine them as the total random strings for the whole experimental data. Let $nH_\infty(A|X) = k_1$ and $nH_\infty(B|XAY) = k_2$, from the definition of block-wise source [175], the output A and B form a block-wise source. Thus we can use two extractors to extract the randomness in A and B. First, we use one Toeplitz hashing matrix to extract the k_2 bits randomness from Bob's side, and then we construct another Toeplitz hashing matrix to extract the k_1 bits randomness from Alice's side. Then the final output string is $n(k_1 + k_2)$ bits.

Since we use two extractors to get the final output string m, Eqn. 3.7 is altered to

$$\Delta_i = \frac{1}{2}\sqrt{2^{m_i - k_i}}, \tag{6.17}$$

where $i \in \{1, 2\}$, and k_1 is the total min-entropy of SDI randomness in Alice's experimental data, and k_2 is the left independent SDI randomness in Bob's side. With two extractors, the distance between m and a uniform random distribution is $\Delta = \Delta_1 + \Delta_2$ [175]. And m is written as

$$\begin{aligned} m &= m_1 + m_2 \\ &= \lfloor k_1 + \log_2 (2\Delta_1)^2 \rfloor + \lfloor k_2 + \log_2 (2\Delta_2)^2 \rfloor \\ &= \lfloor k_1 - 2\log_2 \frac{1}{2\Delta_1} \rfloor + \lfloor k_2 - 2\log_2 \frac{1}{2\Delta_2} \rfloor, \end{aligned} \tag{6.18}$$

The total SDI randomness in Eqn. 6.16 can be quantified by the RSP-dimension witness model. From the formula in [4], the SDI randomness k_2 on Bob's side can be calculated, then the SDI randomness k_1 can be derived directly. With the SDI randomness from both sides, we can extract them by two Toeplitz hashing matrices.

From Eqn. 6.17, m_1 and m_2 can be represented as

$$m_i = k_i - 2 \log_2 \frac{1}{2\Delta_i} \tag{6.19}$$

where $i \in \{1, 2\}$, and usually we choose $\Delta_1 = \Delta_2 = \Delta$. Then with m_1 and m_2 derived, for Alice's side, we use $n + m_1 - 1$ bits to construct a $n \times m_1$ Toeplitz matrix (where n is the length of input raw random strings), and next use the original output string a of Alice, which is $1 \times n$, afterward, we get a m_1 bits long almost uniformly distributed random sequence with at most Δ deviation from a uniformly distributed random sequence. Similarly, the total min-entropy m_2 in Bob's raw output string can be extracted. Combining m_1 and m_2, the final SDI randomness of the joint outcome events can be obtained.

6.4.3 Experimental data analysis

Using the analytical model in [82] and taking the confidence level as 0.99 [82, 4], the min-entropy in our Bell test data can be quantified. For the $|\Psi^+\rangle$ state, the data resulted in $S = 2.085$, and with a total amount of events $n = 27,885$, considering the 99% confidence level, no DI randomness can be quantified for this entangled state. Performing the same task for the 27,683 events from the $|\Psi^-\rangle$ state, the value of S is 2.177. The min-entropy of the DI randomness is 531 bits, much smaller than our SDI randomness, as shown in the following text.

Next, we apply the SDI model to bound the min-entropy of the randomness produced in the Bell test [62] and then extract the randomness with hashing functions. The confidence level is taken as 99%, and the hashing error is chosen as 0.001. We use the Toeplitz-hashing extractors mentioned in Chapter 3 to extract the bounded randomness. Considering the $|\Psi^+\rangle$ state, the Bell test data resulted in $S = 2.085$, with a total number of events $n = 27,885$. We calculate the RSP-dimension witness value for this entangled state as $W_{\text{rsp}} = 0.542$. Using the Toeplitz matrices twice, as suggested in the above subsection, the SDI randomness extracted in all events amounts to $3,821$ bits, which is a significant improvement compared to the 0 bits in the DI model of [82].

Performing the same task for the $27,683$ events from the $|\Psi^-\rangle$ state, the value of S is 2.177, and the RSP-dimension witness value is $W_{\text{rsp}} = 0.591$. The extracted SDI randomness amounts to $4,660$ bits, which is much larger than 531 bits DI randomness [3].

The NIST Statistical Test Suite results for the extracted bits of $|\Psi^-\rangle$ state are shown in Fig (6.4). Although this extracted random sequence does not pass all the tests, it still passes 13 out of 16 tests. The failure of the three tests is due to the insufficient length of the random sequence.

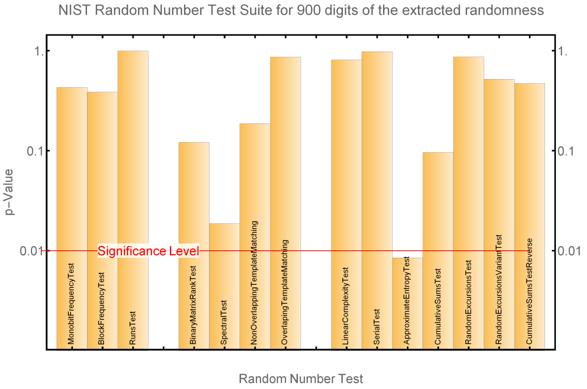


Figure 6.4: NIST Statistical Test Suite results. The test is done for the SDI randomness for $|\Psi^-\rangle$ state. The final random string does not pass all the tests due to its insufficient length.

We also draw the binary image in Fig 6.5 of the extracted random bits, and no predictable patterns are shown in this figure.

[3] The 531 bits randomness is the bounded randomness from the DI model, not considering the hashing error. If we consider a 0.001 hashing error, this value will decrease to 495 bits.

Figure 6.5: Binary image of the extracted random bits from $|\Psi^-\rangle$. In this binary image, no obvious patterns can be seen.

6.5 Conclusion

This chapter presents two methods to bound the min-entropy of the quantum randomness in our Bell test data. The DI model from [82] is based on Bell's theorem, and its applicability holds especially for the CHSH-variant of the test [73]. For all the $55,568$ events, the min-entropy of the DI randomness is 531 bits, which amounts to 0.956×10^{-2} bits per event. An extended RSP-dimension witness model is newly designed for the same Bell test. For all $55,568$ events data, the total min-entropy of the extracted SDI randomness is $8,481$ bits, corresponding to 0.153 bits per event, which is significantly higher than 0.956×10^{-2} bits per event data.

Our RSP-dimension witness model tremendously improves the bound of the randomness from the Bell test data without using Bell's theorem. This model can still certify quantum randomness when the Bell inequality is not violated. Although the SDI model offers relatively weaker security guarantees for randomness than the DI model, it still provides certified randomness. Additionally, the SDI model's requirements can be met using standard technologies, which are much less complex than those required for a loophole-free Bell test. This is one important step towards the practical use of the Bell test in randomness generation.

7 Randomness from a nuclear spin system

The previous chapter showed how to turn the Bell test into a prepare-and-measure experiment scheme. Furthermore, within this prepare-and-measure scheme, we can use the dimension witness model [83] to quantify the entropy of randomness in the raw experimental data. In this chapter, we continue the discussion of QRNG based on dimension witness. Here we present a proof-of-concept random number generator based on a nuclear spin state system instead of a photonic system discussed in [4]. Compared with the original dimension witness protocol in [4], our work extends the dimension witness protocol from dimension 2 to dimension 3, and our QRNG is based on a nuclear spin system instead of a photonic system.

Before the discussion of our protocols, we first introduce the nuclear spin state in the NV center.

7.1 Nuclear spin state in NV color center

The nuclear spin state in our experiment belongs to the nucleus of the nitrogen atom of a single NV defect in a diamond. The NV defect, also known as the NV color center, is a very important physical system for quantum technologies, including quantum sensing, information processing, and communications. The nuclear spin system inside the NV color center is well isolated from the environment and can be operated at room temperature. Their creation and control are relatively easy [90, 176, 91]. This stable quantum system has been used in the quantum computing area for years [92, 91, 93]. In the previous chapter, we showed how to use the NV color center as a single-photon source to generate quantum randomness. This chapter demonstrates how to employ the nuclear spin state within the NV center to generate quantum randomness.

The schematic structure of the NV center is shown in Fig. 7.1. In the NV center, a single NV defect in diamond is coupled to a single nuclear spin of ^{13}C and ^{14}N. The defect consists of a nitrogen impurity next to a vacancy in the diamond lattice. The method to prepare such a NV center includes chemical vapor deposition (CVD) diamond synthesis process [177], radiation damage and annealing [115], ion implantation and annealing in bulk and nanocrystalline diamond [178]. There are two different charge states in an NV center, including the negative charge state (NV^-) and neutral charge state (NV^0). The optical zero photon lines (ZPLs) of NV^- and NV^0 are shown in Fig. 7.2.

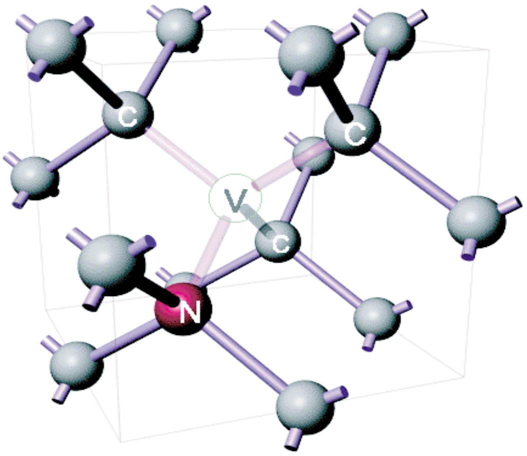


Figure 7.1: Schematic structure of the NV center. The figure is adapted from [114]

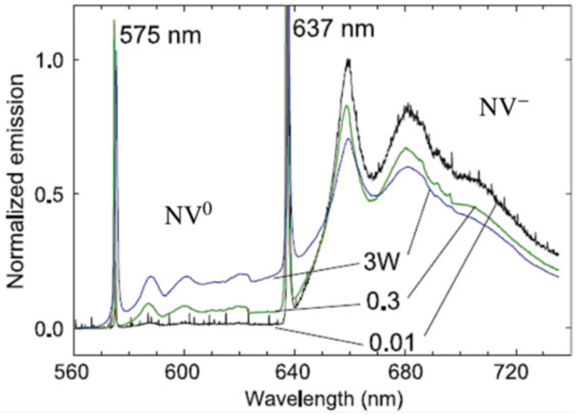


Figure 7.2: Normalised emission spectra of NV^- and NV^0. These spectra are obtained at low temperature(10K) for different excitation powers. The figure is adapted from [115]

In this thesis, we focus on the negative charge state NV^- due to its increased stability at room temperature and its success in various applications [179, 180, 181, 182].

For NV^-, an additional electron is trapped in the vacancy. The nuclear spin states of nearby atoms like ^{13}C [90] and ^{14}N [183] are coupled with the electron spin. The electron spin coherently couples with the nuclear spins. By tuning the external magnetic field, the NV center can be coupled with a desired nearby nucleus, for example, coupling with ^{14}N. The electron-nuclear spin system can be effectively controlled with a specific microwave pulse. With single-shot readout [184, 183], repetitive optical readout of the electron spin will reveal its coupled nuclear spin state due to the coupling between the electron and nuclear spin systems.

The basic idea of single-shot readout is shown in Fig. 7.3. The coupling scheme of the electron-nuclear spin state is shown in Fig. 7.3A, where Ψ_n represents the spin state of the nucleus(^{14}N here). This figure shows that the electron-nuclear system is equivalent to a controlled not (CNOT) system, which is possible because of the long coherence time of the NV center. Fig. 7.3B shows the real-time dynamics of a single nuclear spin, and we can see the abrupt, discontinuous evolution of the nuclear spin state. The fluorescence intensity in Fig. 7.3B is obtained through single-shot readout. In this process, the electron spin is optically pumped into one sublevel $|0_{\mathrm{e}}\rangle$(corresponds to $m_{\mathrm{S}} = 0$) of its triplet ground state($S = 1$), meanwhile, the nuclear spin of ^{14}N is in an incoherent mixture of the its eigenstates $|-1_{\mathrm{I}}\rangle$, $|0_{\mathrm{I}}\rangle$, and $|+1_{\mathrm{I}}\rangle$ ($m_{\mathrm{I}} = -1, 0, +1$). Then the application of a narrowband, nuclear-spin state-selective microwave (MV) π pulse will flip the electron spin into $|-1_{\mathrm{e}}\rangle$ conditioned on the state of the nuclear spin state.

Because the fluorescence intensity between electron spin states $|0_{\mathrm{e}}\rangle$ and $|-1_{\mathrm{e}}\rangle$ differs by roughly a factor of 2, the two spin states can be distinguished by shining a short laser pulse, which will destroy the electron spin state but leave the corresponding nuclear spin state almost undisturbed under the experimental conditions. So, repeated measurement of electron spin states allows a nondestructive accumulation of the fluorescence signal, which will optically determine the nuclear spin state.

The nuclear spin state can be initialized into the superposition of two states (for example, $|0_{\mathrm{I}}\rangle$ and $|-1_{\mathrm{I}}\rangle$), this will formulate a qubit state. The nuclear spin state system can be projected into one of the two eigenstates without demolishing it by using the single-shot readout method mentioned above. The projection will be random from the perspective of quantum mechanics. Since single-shot readout is a quantum nondemolition measurement, the same nuclear spin state can be prepared and measured consecutively, and a random sequence can be generated by detecting different eigenstates.

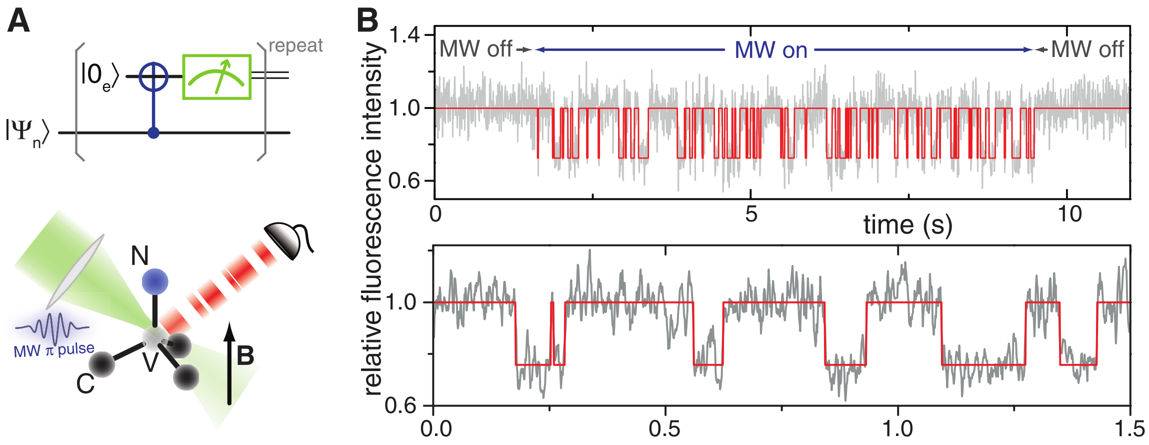


Figure 7.3: quantum jumps of a single nuclear spin in real-time. Single-shot readout reveals quantum jumps of a single nuclear spin in real time. Figure is adapted from [183]

Next, the experimental setup and the generation of raw bits are described. The experiment was mainly performed by my colleagues Minsik Kwon and Dr. Vadim Vorobyov. This work is also summarized in our manuscript, which is currently in preparation.

7.2 Experimental setup and generation of raw bits

Our QRNG protocol is based on dimension witness, where the prepare-and-measure (P&M) scenario is important. The P&M scheme is already explained in Chapter. 2. In our experiment here, the prepared states ρ_x now have six possible states: $|0\rangle, |1\rangle, |+\rangle, |-\rangle, \frac{|0\rangle+\mathbf{i}|1\rangle}{\sqrt{2}}, \frac{|0\rangle-\mathbf{i}|1\rangle}{\sqrt{2}}$ and the measurement y now has three different measurements $\sigma_1, \sigma_2, \sigma_3$. The P&M scenario is slightly altered in this experiment: Alice and Bob are in the same location, so Bob measures the states immediately after Alice prepares them.

As mentioned above, the nuclear spin states used in our protocol are from the nucleus of a nitrogen atom [^{14}N] of a negatively charged NV center. The energy level of the NV$^-$ and a sketch of our experimental setup is shown in Fig. 7.4.

First, we use a pulse of a 520 nm green laser to initialize the NV center into the negative charge state NV$^-$, and the electron spin state into $|0_\mathrm{e}\rangle$ (corresponds to $m_\mathrm{S} = 0$). In the meantime, the nuclear spin state is in an incoherent mixture of the its eigenstates $|-1_\mathrm{I}\rangle$, $|0_\mathrm{I}\rangle$, and $|+1_\mathrm{I}\rangle$ ($m_\mathrm{I} = -1, 0, +1$). Then we apply a series of microwave (MW) π pulses and radio-frequency (RF) π pulses to initialize the nuclear spin states from $|0_\mathrm{I}\rangle$, $|+1_\mathrm{I}\rangle$ into $|-1_\mathrm{I}\rangle$. The charge state preparation fidelity and the nuclear spin state initialization fidelity can be probed

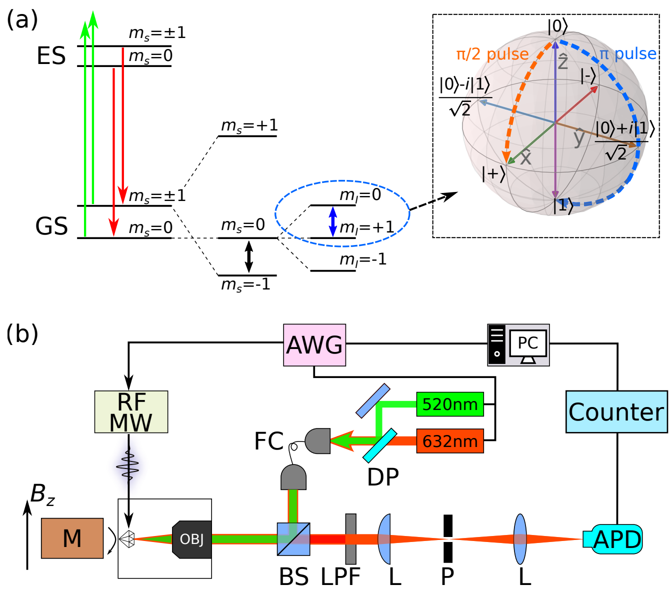


Figure 7.4: Experimental scheme to prepare and measure nuclear spin states. (a) Energy level of electron spin state and nuclear spin state of the NV$^-$ state, where **ES** means excited state and **GS** means ground state. m_s is the electron spin state, and m_I is the nuclear spin state. We choose the nuclear spin state $|0_\mathrm{I}\rangle$ and $|+1_\mathrm{I}\rangle$ to prepare our states, which is shown in a Bloch-sphere in the right. **(b)** The sketch of our experimental setup. In the figure **M** represents the permanent magnet, **RF**: radio-frequency pulse, **MW**: microwave pulse, **OBJ**: objective, **FC**: fiber coupler, **AWG**: arbitrary waveform generator, **DP**: dichroic polarizer, **BS**: beam-splitter, **LPF**: long pass filter, **L**: lens, **P**: pinhole, **APD**: avalanche photodiode, and **Counter** here is a time tagger. We use a 532 nm laser (632 nm laser) for charge state NV^0 (NV$^-$) readout.

with single-shot readout. Their single-shot readout results are shown in the photon counting histograms of Fig. 7.5(b). In the left sub-figure, different distinguishable peaks corresponding to different NV charge states are shown. Similarly, in the middle sub-figure, different nuclear spin states correspond to different peaks of this histogram. By setting the green threshold as shown in the figure, the nuclear spin state $|-1_\mathrm{I}\rangle$ (fluorescence below threshold) can be distinguished from the other nuclear spin states (fluorescence above threshold). The fidelity of charge state NV^{-1} is $F_1 = 1 - \epsilon_1 = 98.16 \pm 0.36\%$, and the fidelity of nuclear spin state initialization is $F_2 = 1 - \epsilon_2 = 95.17 \pm 0.40\%$, where $\epsilon_{1,2}$ are the corresponding errors when distinguishing the states in the two histograms with the given thresholds in Fig. 7.5(b) (left and middle).

After the initialization stage, the nuclear spin state is known, and it is in state $|-1_\mathrm{I}\rangle$. With this known nuclear spin state, we can start our P&M protocol. In this protocol, six possible states need to be prepared, and they can be represented in a Bloch sphere as shown in Fig. 7.4. Thus, these six states can be prepared within a qubit system. In this experiment here, we choose nuclear spin state $|0_\mathrm{I}\rangle$ and $|+1_\mathrm{I}\rangle$ to form the two-level qubit system.

States	RF operations	Nuclear spin state
$\|0\rangle$	$R_{y_{\text{axis}}}^{-1,0}(\pi)$	$\|0_{\text{I}}\rangle$
$\|1\rangle$	$R_{y_{\text{axis}}}^{-1,0}(\pi)$+$R_{y_{\text{axis}}}^{0,+1}(\pi)$	$\|1_{\text{I}}\rangle$
$\|+\rangle$	$R_{y_{\text{axis}}}^{-1,0}(\pi)$+$R_{y_{\text{axis}}}^{0,+1}(\pi/2)$	$\frac{\|0_{\text{I}}\rangle+\|1_{\text{I}}\rangle}{2}$
$\|-\rangle$	$R_{y_{\text{axis}}}^{-1,0}(\pi)$+$R_{y_{\text{axis}}}^{0,+1}(-\pi/2)$	$\frac{\|0_{\text{I}}\rangle-\|1_{\text{I}}\rangle}{2}$
$\frac{\|0\rangle+\mathbf{i}\|1\rangle}{\sqrt{2}}$	$R_{y_{\text{axis}}}^{-1,0}(\pi)$+$R_{x_{\text{axis}}}^{0,+1}(-\pi/2)$	$\frac{\|0_{\text{I}}\rangle+\mathbf{i}\|1_{\text{I}}\rangle}{2}$
$\frac{\|0\rangle-\mathbf{i}\|1\rangle}{\sqrt{2}}$	$R_{y_{\text{axis}}}^{-1,0}(\pi)$+$R_{x_{\text{axis}}}^{0,+1}(\pi/2)$	$\frac{\|0_{\text{I}}\rangle-\mathbf{i}\|1_{\text{I}}\rangle}{2}$

Table 7.1: State preparation stage, how to prepare different states with different RF operations from the same initialized nuclear spin state $|-1_{\text{I}}\rangle$.

Since the nuclear spin state is initialized into $|-1_{\text{I}}\rangle$ state, in order to prepare the six states within state $|0_{\text{I}}\rangle$ and $|+1_{\text{I}}\rangle$, a series of RF pulses need to be applied to flip the nuclear spin state from $|-1_{\text{I}}\rangle$ into the qubit system. For example, to prepare state $|0\rangle$, we need to apply a RF π pulse between $|-1_{\text{I}}\rangle$ and $|0_{\text{I}}\rangle$ ($R_{y_{\text{axis}}}^{-1,0}(\pi)$) to flip the spin state from $|-1_{\text{I}}\rangle$ into $|0_{\text{I}}\rangle$, which corresponds to state $|0\rangle$. If we want to prepare another state, additional RF pulse $R_i^{0,+1}(\theta)$ along the axis i ($i \in \{x_{\text{axis}}, y_{\text{axis}}, z_{\text{axis}}\}$) can be applied. Table. 7.1 shows the states and their corresponding RF pulse operations. For instance, by applying a RF pulse $\pi/2$ along the y axis ($R_{y_{\text{axis}}}^{0,+1}(\pi/2)$), the nuclear spin will be in state $(|0_{\text{I}}\rangle + |1_{\text{I}}\rangle)/\sqrt{2}$, which corresponds to state $|+\rangle$.

Following the state preparation, we measure the prepared nuclear spin state. Take state $|+\rangle$ as an example, it is a superposition state of $|0_{\text{I}}\rangle$ and $|+1_{\text{I}}\rangle$, when we measure it in basis $\hat{z}$, the state $|+\rangle$ will collapse into state $|0_{\text{I}}\rangle$ and $|+1_{\text{I}}\rangle$ randomly. The single-shot readout results of the collapsed nuclear spin states are shown in Fig. 7.5(b) (right subfigure), and we assign the left peak as random bit “1," and the right peak as “0". The distribution of raw bits is not perfectly separated from each other. A small fraction of them is still overlapped. In order to distinguish the results unambiguously, as shown in Fig. 7.6, a threshold is set to identify them explicitly. The distinguish fidelity corresponds to this threshold is $F_3 = 1 - \frac{\epsilon_l + \epsilon_r}{2} = 98.55 \pm 0.46\%$, where $\epsilon_{l,r}$ are the errors in each histogram when applying the green dashed line as the threshold to distinguish bits “1" and “0" in Fig. 7.6.

In Fig. 7.6 (a), the threshold is fixed for the measurement results of all 18 different P&M pairs (six preparations and three measurements, in total, there are 18 different P&M pairs) in the whole experimental data. With the threshold being settled, the probability distribution of each P&M combination and their corresponding error bars are shown in Fig. 7.6(b).

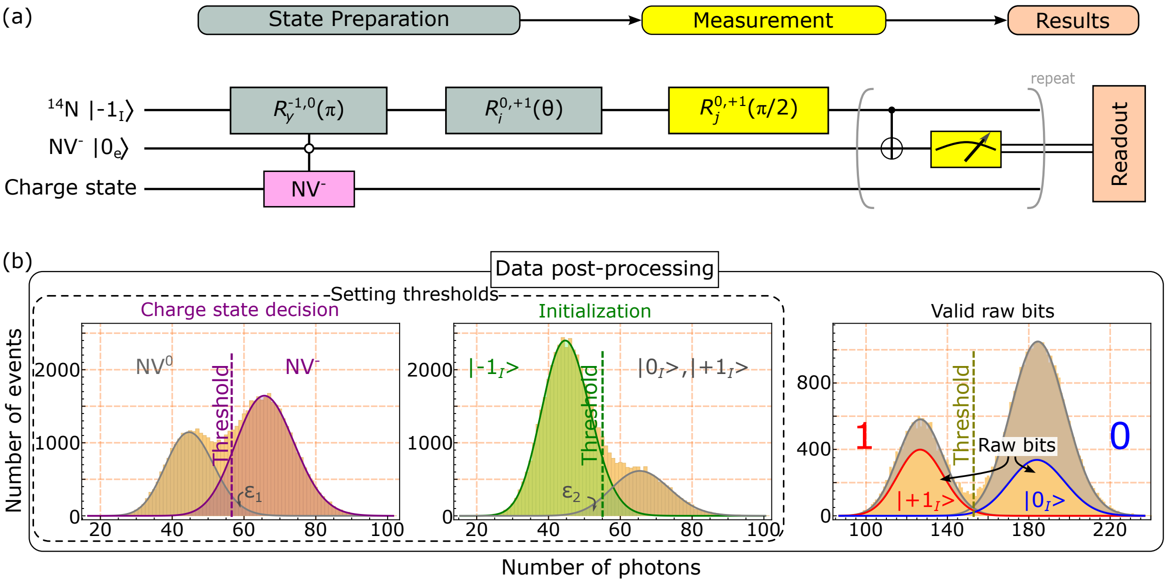


Figure 7.5: The P&M scenario and the data post-processing. All the histograms in this figure are fitted by Gaussian distributions (solid lines). **(a)** Quantum circuit representation of our P&M scenario. First, we pump the NV color center into NV$^-$ and initialize the nuclear spin state of ^{14}N into $|-1_\mathrm{I}\rangle$. Then in the state preparation step, we apply a π pulse to this nuclear spin state and rotate it into $|m_I = 0\rangle$, which corresponds to the state $|0\rangle$. With further π or $\pi/2$ pulses, the other five states $|1\rangle, |+\rangle, |-\rangle, \frac{|0\rangle+\mathbf{i}|1\rangle}{\sqrt{2}}$, and $\frac{|0\rangle-\mathbf{i}|1\rangle}{\sqrt{2}}$ can be prepared. This is represented as $R_i^{0,+1}(\theta)$ ($\theta = \{\pi, \pi/2\}$) in the figure. In the measurement step, we apply a $\pi/2$ pulse ($R_j^{0,+1}(\pi/2)$) along the $\hat{x}, \hat{y}$ or $\hat{z}$ axis to perform a corresponding measurement, then single-shot readout is used to get the state of the nuclear spin system. **(b)** In the post-processing process, we set thresholds for the charged state and nuclear spin state initialization distribution histograms. For the charge state single-shot readout histogram, the threshold is set as 56, and for the photon counts from NV$^-$ state is larger than this value. In the initialization histogram, the threshold is set as 55, and the spin state $|-1_\mathrm{I}\rangle$ corresponds to the photon count results, which are no larger than this value. ϵ_1 and ϵ_2 are the corresponding errors of choosing the given thresholds in the histograms. In the valid raw bits histogram, the threshold is 153, we assign the left peak as raw bits “1" and the right peak as “0".

After fixing the thresholds for charge state decision, nuclear spin state initialization, and the raw bits, the total fidelity of each P&M pair can be calculated by the following equation

$$F_{\text{total}} = F_1 \cdot F_2 \cdot F_3 = 92.06 \pm 0.67\% \tag{7.1}$$

With the raw bits from our experimental setup, next, we develop two dimension witness [83] protocols to quantify the entropy in the raw bits.

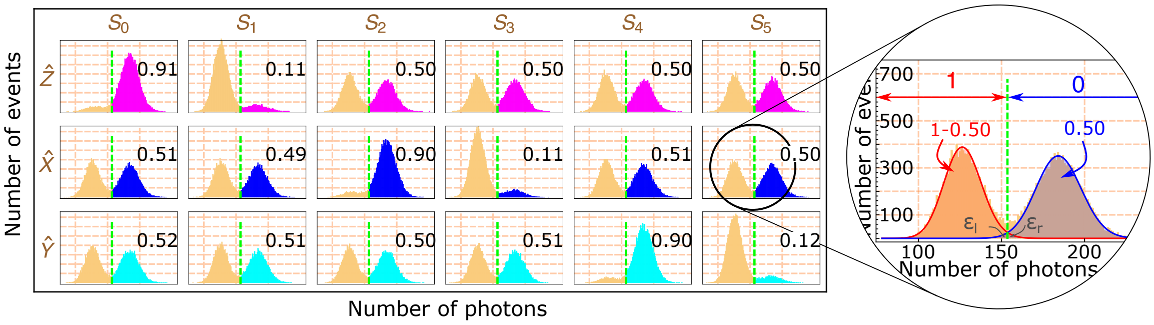


Figure 7.6: Experimental results analysis. Photon-counting distribution of 18 different P&M pairs. The vertical axis of all 18 sub-figures is the number of events for different photon counts, and the horizontal axis is the number of photon counts. The dashed green line is the threshold line that is used to differentiate the raw bits “1" and “0". The photon counts value of the green dashed line here is 153, and in the appendix, we show how to get this threshold in detail. In the right big circle of this sub-figure, the photon counts distribution of one specific P&M ($S_5 - \hat{x}$) is shown. The probability of getting “1" is calculated by the photon counts, which are smaller than the threshold line, and “0" is the photon counts, which are larger than the threshold line. $\epsilon_{l,r}$ are the measurement errors of each peak when applying the green dashed line as the threshold.

7.3 Randomness certification protocols

Before we proceed with the dimension witness protocols, we describe the assumptions needed for our two protocols. The assumptions mentioned here are similar to the assumptions in [4, 3], but for the self-consistence of this chapter, we rephrase them here. Our two protocols are based on the same experimental device, and they require the same assumptions: (1) the state preparation and measurement settings (x, y) are independently chosen; (2) the preparation and measurement devices are independent of each other; (3) the information in the measurement results of each side is contained in a two-dimensional quantum subspace; (4) the system is memoryless, and subsequent outcomes are not directly correlated.

Assumptions (1) and (2) require that the state preparation and measurement devices are independent of each other. Assumption (1) also implies that the devices that are used to generate the input strings x, y are not correlated with the measurement devices. Subsequently, x, y can be pseudo-random numbers as long as they are independent of each other and the measurement apparatus. These two assumptions can be easily realized if the experimental devices are not manufactured by malicious producers. In our case, the state preparation and measurement devices are error-prone. However, we assume their errors are independent of each other, and there are no pre-established correlations between the two devices. Assumption (3) means that a qubit can describe the measurement results of each measurement setting. About assumption (4), it requires the previous measurement result of the device does not affect the measurement results of the following measurements. The experimentally relevant afterpulsing effect can likely weaken this assumption. But still, randomness can be bounded in the presence of certain memory effects since the memory effect of the device can be explained by shot noise.

In order to analyze the memory effect of our experimental setup, we calculate the differences in conditional probabilities of the adjacent P&M runs. There are four conditional probabilities $p(1|0),\ p(1|1),\ p(0|0),\ p(0|1)$. Among them, $p(1|0)$ means the probability of getting results “1" in the current P&M run, with result “0" from the previous run. The other three conditional probabilities have equivalent definitions. Ideally, all these conditional probabilities should be 0.5, meaning the previous measurement result does not affect the next measurement result. In other words, $p(1|0)-p(1|1)=p(0|1)-p(0|0)=0$. So we can use the difference between $p(1|0)$ and $p(1|1)$ or $p(0|1)$ and $p(0|0)$ to indicate the memory effect of our experimental setup. Since $p(1|0)-p(1|1)=p(0|1)-p(0|0)$ holds true in general[1], we only use the difference $p(1|0)-p(1|1)$ in the following calculation. Considering the i-th and the $(i+1)$-th P&M run, there will be a joint event $\{b_i, x_i, y_i, b_{i+1}, x_{i+1}, y_{i+1}\}$. We want to show that for any $\{x_i, y_i\}$, the result b_i does not affect the distribution of b_{i+1}. This can be represented as the differences between the conditional probabilities $p(1_{i+1}|0_i)$ and $p(1_{i+1}|1_i)$. In Fig. 7.7, such differences from each i-th P&M pair are shown. As expected, all the conditional probability differences can be explained by the 3σ shot noise, which indicates that the memory effect of our experimental device is limited, and we can use dimension witness to quantify the entropy of randomness in our nuclear spin state quantum randomness generation system.

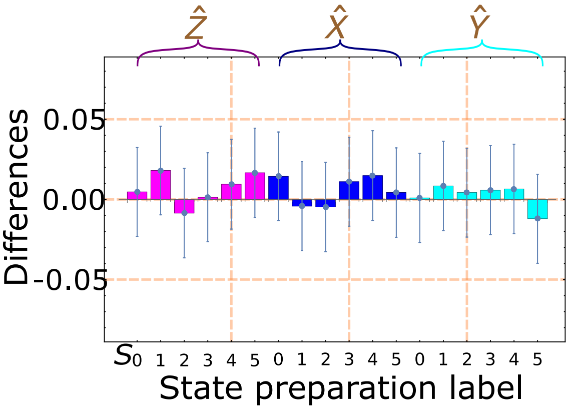


Figure 7.7: Conditional probability differences of different P&M pairs. This figure shows that the conditional probability differences are not zero, but they can all be explained by the 3σ shot noise. In other words, our experimental setup can be considered memoryless.

[1] This is straightforward considering the fact that $p(1|0)+p(0|0)=p(0|1)+p(1|1)=1$.

7.3.1 W_2 Model description

In the model of [4], the P&M scenario is done in a two-dimensional Hilbert space (i.e., a plane in a Bloch sphere). Here we extend this original model into a three-dimensional Hilbert space, more specifically, three mutually orthogonal planes. The state ρ_x, which is prepared by Alice, is randomly chosen from six possible preparations : $\vec{S}_0 = -\vec{S}_1 = \hat{z}$, $\vec{S}_2 = -\vec{S}_3 = \hat{x}$ and $\vec{S}_4 = -\vec{S}_5 = \hat{y}$. After the preparation, Bob measures the states in measurements : $\vec{T}_0 = \hat{z}$, $\vec{T}_1 = \hat{x}$ and $\vec{T}_2 = \hat{y}$, and get a binary result b.

In the experimental scheme above, the state preparations and measurements can be categorized into three planes, and a two-dimensional dimension witness W_2 in Eqn. 2.12 can be constructed in each plane. For instance, in the Bloch sphere of Fig. 7.4(a), within $\hat{x}\hat{z}-$plane, the states $|0\rangle, |1\rangle, |+\rangle, |-\rangle$, measurement bases $\hat{x}, \hat{z}$, and the measurement results can be used to construct a 2-dimensional dimension witness $W_{\hat{\mathrm{x}}\hat{\mathrm{z}}}$. The dimension witness $W_{\hat{\mathrm{x}}\hat{\mathrm{y}}}$ in the $\hat{x}\hat{y}-$plane and the dimension witness $W_{\hat{\mathrm{y}}\hat{\mathrm{z}}}$ in the $\hat{y}\hat{z}-$plane can be constructed in a similar fashion.

Since three 2-dimensional dimension witnesses can be constructed in different planes, we use the dimension witness formula in [4] three times. In an ideal condition, with the preparations and measurements mentioned above, in each use of the formula, the maximal value of W_2 can be reached, and we can bound randomness with min-entropy $H_{\min} = 0.2284$. Considering we only utilize $\frac{4}{9}$ raw bits data [2] in each application of the protocol, so the total output entropy is $3 \times \frac{4}{9} \times 0.2284 = 0.3045$ bits per event in an ideal case.

The limitation of this protocol is obvious. It is a repeated usage of the W_2 protocol in [4]. It means that the protocol can only be applied to the specific states and measurements which are distributed on the intersection lines of three different planes in a Bloch sphere so the W_2 dimension witness can be constructed in each plane. In our experiment scheme, the states and measurements are on the intersection lines of $\hat{x}\hat{y}-$, $\hat{x}\hat{z}-$, and $\hat{y}\hat{z}-$ planes. If the six preparations S_x and three measurements T_y do not form three different planes, this protocol will be inapplicable. For example, consider rotating $|0\rangle$ and $|1\rangle$ by an angle $0 < \theta < \pi/2$ in the $\hat{x}\hat{z}-$plane of the Bloch sphere. After the rotation, dimension witnesses $W_{\hat{\mathrm{x}}\hat{\mathrm{z}}}$ and $W_{\hat{\mathrm{x}}\hat{\mathrm{y}}}$ can still be constructed, but $W_{\hat{\mathrm{y}}\hat{\mathrm{z}}}$ cannot be calculated, since the $\hat{y}\hat{z}-$plane does not contain four

[2] We have six states in the state preparation stage and three different measurement settings. As we choose states and measurements randomly, in total, there will be 18 different combinations appearing with the same probability. In each construction of the W_2, we choose four states and two measurements, which have eight different combinations. So, on average, in one event, we have $\frac{4}{9}$ fraction of data.

states anymore as shown in Fig. 7.8. Next, we develop a more robust protocol that utilizes a three-dimension dimension witness to quantify the min-entropy of the randomness.

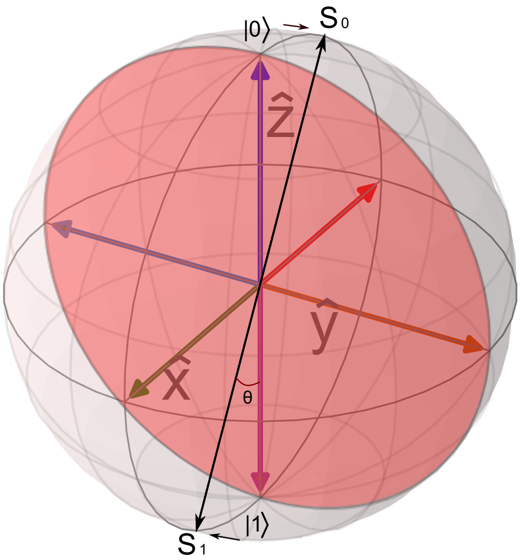


Figure 7.8: One failure scenario of the W_2 model. By rotating $|0\rangle$ and $|1\rangle$ by an angel $0 < \theta < \pi/2$ in the $\hat{x}\hat{z}-$plane, they will be outside the $\hat{y}\hat{z}-$plane, and the dimension witness $W_{\hat{y}\hat{z}}$ cannot be constructed.

7.3.2 W_3 Model description

There is another way to bound more randomness from the same data. The first protocol is based on the W_2 protocol [4], and in an ideal case, the entropy in the output randomness is 0.3045 bits per raw bit. In our second protocol, which is our main work in this chapter, we utilize a 3-dimensional dimension witness W_3 (defined below) to quantify the entropy of the randomness, and the entropy increases to 0.3425 bits per raw bit.

In order to construct W_3, we need 6 preparation states and 3 measurement bases as proved in [83]. The six states in preparation need to be in a dimension d which is no smaller than $\lceil\sqrt{3}\rceil$, so we can get a nonzero W_3 and then we can quantify the quantumness involved in the measurement process by W_3.

Our experiment has six input states ρ_x and three measurement bases. The Hilbert space of each state has a dimension $d = 2$, which is equal to $\lceil\sqrt{3}\rceil$. Therefore, a three-dimensional dimension witness W_3 can be constructed for our experimental scheme. This W_3 is defined as

$$W_3 = \begin{vmatrix} p(0,0) - p(1,0) & p(2,0) - p(3,0) & p(4,0) - p(5,0) \\ p(0,1) - p(1,1) & p(2,1) - p(3,1) & p(4,1) - p(5,1) \\ p(0,2) - p(1,2) & p(2,2) - p(3,2) & p(4,2) - p(5,2) \end{vmatrix} \tag{7.2}$$

where $p(x,y)$ in the equation is $p(b = 0 \mid x, y)$. When $0 < W_3 \leq 1$, quantumness is involved in the measurement data. This quantumness means the measurement results are not deterministic; thus, we can get quantum randomness from the measurement results.

The entropy of the randomness per raw bit is quantified by the conditional min-entropy $H_\infty(B|XY)$ (where B, X, Y represent the sets of random variables b, x, y). It is defined as $H_\infty(B|XY) = -\log_2 p_{\text{guess}}$. Assuming uniformly distributed input x and y, the guessing probability p_{guess} is derived as

$$\begin{aligned} p_{\text{guess}} &= \frac{1}{18} \sum_{a,b} \max_m p(m|a,b) \\ &\leq \frac{1}{3} \max_a \sum_b \max_m p(m|a,b) \\ &\leq \frac{1}{3} \left(\frac{3 + \sqrt{3}\sqrt{1 + 2\cos\theta}}{2} \right) \\ \theta &= \cos^{-1}\left(\frac{1}{2} \sqrt{3 + 2\sqrt{1 + 8W_3^2} \cos\left(\frac{1}{3} \left(\pi + \tan^{-1}\left(1 - 20W_3^2 - 8W_3^2, 8W_3 \sqrt{(1 - W_3^2)^3} \right) \right) \right)} \right) \end{aligned} \tag{7.3}$$

Suppose the preparation states and measurement settings are the same as the previous model. Then in each round of the protocol, we choose settings among the six possible preparations $x = \{0, 1, 2, 3, 4, 5\}$, and three measurement bases $y = \{0, 1, 2\}$, resulting in a binary outcome $b = \{0, 1\}$. The distribution of preparations a and measurements b are uniform, then $W_3 = 1$ can be reached. In this case, the min-entropy is $H_\infty(B|XY) = -\log_2 p_{\text{guess}} = 0.34$ bits per raw bit. When the system is in a non-ideal situation, such as misalignment, we will get $0 < W_3 < 1$. However, the guessing probability in Eqn. 7.3 is still larger than 0.5, and quantum randomness can be certified.

Both W_2 and W_3 models are based on the dimension witness [83], and their confidence level analysis is similar to the one in chapter 6.

7.4 Results analysis

The same QRNG scheme was performed three times in our nuclear spin system. Only the data set from the first implementation is analyzed here. The analysis for the remaining data sets is similar, and their results are presented in the appendix.

The experimental running time is about 7 hours, and the data set contains 418,666 raw bits. In the W_2 model, with a confidence level of 99% and hashing error 0.001 [145], the entropy of randomness per raw bit is 0.0541 bits. The total extractable randomness is 23,525 bits, and the randomness generation speed is about 0.87 bits per second. While in our W_3 model, with the same confidence level and hashing error, the entropy of randomness per raw bit is 0.0825 bits. The total extractable random bits is 35,632 bits, and the randomness generation speed is about 1.33 bits per second.

In Fig. 7.9, the theoretical curves of our W_3 and W_2 models are shown. From the above analysis, we can see that in our experimental data, compared with the W_2 model in [4], 53% more random bits can be certified by our W_3 model.

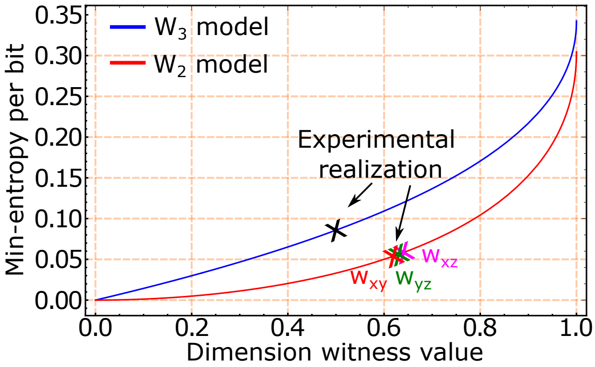


Figure 7.9: Output randomness per raw bit by utilizing different dimension witness protocols. The blue curve (the upper curve) is the theoretical curve of the W_3 model, and the red curve (the lower curve) represents the theoretical curve of the W_2 model. The crosses in blue and red curves are the experimental realization of our two different protocols.

We use the Toeplitz hashing function to do randomness extraction for our W_3 model. We have $n = 418,666$ raw bits, and with a confidence level of 99% and hashing error $\Delta = 0.001$, the length of the extractable random bits is $m = 35,632$ from our W_3 model. The NIST

Statistical Test Suite results are shown in Fig (7.10). We also present the binary image of the extracted random bits by the W_3 model in Fig 7.11.

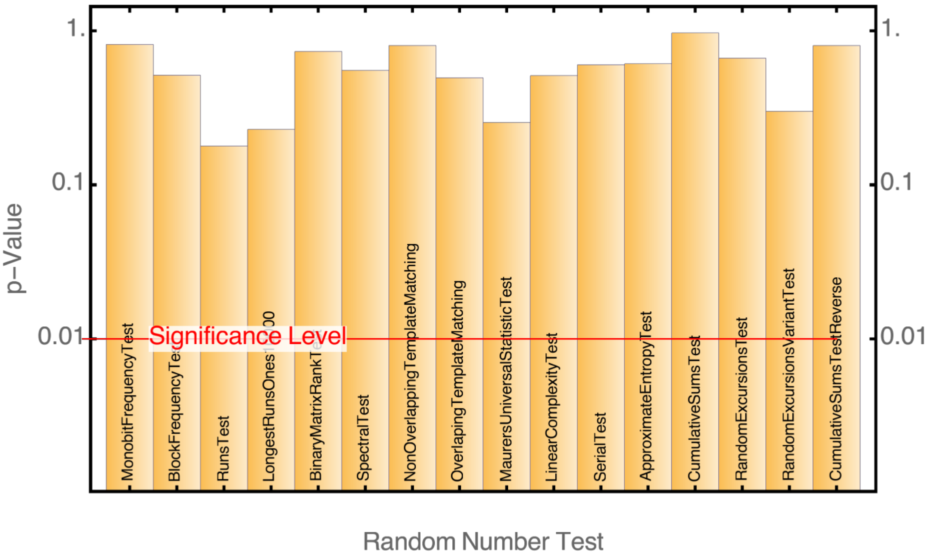


Figure 7.10: NIST Statistical Test Suite results of the extracted random sequence from W_3 model. The extracted random sequence passes all the tests.

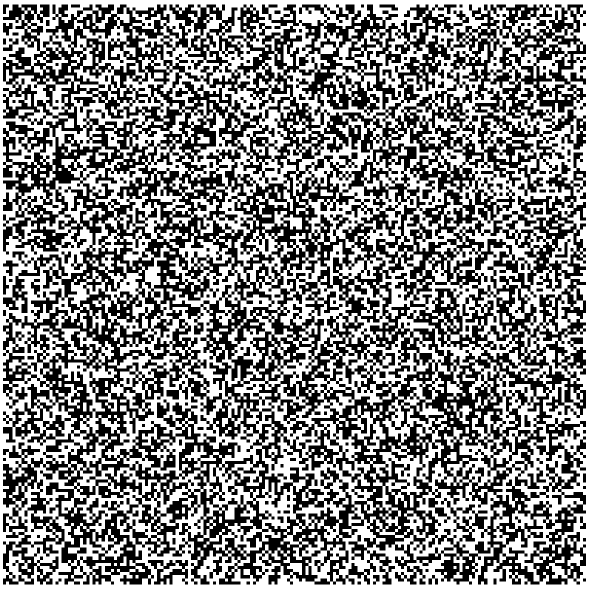


Figure 7.11: Binary image of the extracted random bits from the raw bits. In this binary image from the extracted randomness in W_3 model, no obvious patterns can be seen.

7.5 Conclusion

We generate quantum randomness from the nuclear spin system in a negatively charged single NV center at room temperature and quantify the entropy in the randomness with two different models. Both models are based on the dimension witness [83], which requires no detailed models to describe the experimental devices but only general assumptions, such as the limited dimensionality and the independence of the experimental devices. The first model directly applies the protocol in [4]. In our second dimension witness model, we develop a QRNG protocol based on a three-dimensional dimension witness W_3. We demonstrated how to quantify randomness from our experimental data using the two models.

The first model's application is relatively limited to a particular case where state preparation and measurements are distributed in the intersection lines of three planes. In our case, these three planes are the $\hat{x}\hat{z}-$plane, $\hat{x}\hat{y}-$plane, and $\hat{y}\hat{z}-$plane in a Bloch sphere. The second model is based on a 3-dimensional witness W_3, and it can be applied to any P&M scenario with six input states, three measurement bases, and binary results. Also, from our experimental data, the model based on W_3 can quantify 53% more random bits than the W_2 model. Thus the construction of the W_3 model in our work can certify more quantum randomness in the same experimental data compared with the W_2 model from [4].

8 Conclusion and outlook

8.1 Conclusion

Since the beginning of human civilization, randomness (or random numbers) has been a critical element in various aspects of society, encompassing scientific simulations, information security, and the entertainment industry. True random numbers can only be produced through a quantum mechanical process. However, proving that randomness is derived from a quantum mechanical process rather than classical noise is far from straightforward. In this thesis, we have demonstrated our progress in utilizing the unique characteristics of various quantum phenomena to confirm the presence of quantumness in the generated raw random bits. In essence, we have explored how to certify quantum randomness from raw experimental data, which may contain classical noise.

We introduced standard methods for constructing a QRNG whose randomness can be certified. These methods involve: identifying the quantum phenomenon that can be employed to certify quantum randomness, performing experiments and analyzing the experimental data to reveal quantumness through the identified quantum phenomenon, quantifying the min-entropy of randomness using the quantumness in the data, applying a randomness extractor to obtain a uniformly distributed random sequence, and finally, utilizing the NIST Statistical Test to verify the uniformity of the extracted random sequence.

The four quantum phenomena explored in this thesis are:

- the single-photon antibunching effect,
- wave-particle duality in a delayed-choice experiment,
- non-locality in a loophole-free Bell test, and
- nonzero dimension witness of quantum measurements.

From Chapter 4 to Chapter 7, the QRNG based on each quantum phenomenon is discussed in depth. In each QRNG, the corresponding protocols are constructed by leveraging the quantum nature of the experimental results to quantify the entropy of the raw random bits. Subsequently, Toeplitz matrices as randomness extractors are employed to obtain nearly uniformly distributed random sequences. Finally, the NIST Statistical Test Suite is used to assess the uniformity of the extracted random sequences. Below is a conclusion for each QRNG.

In Chapter 4, the implementation of a single-photon QRNG based on NV center is discussed. Three different QRNG protocols are presented. In the first model, all the raw bits are used to extract randomness, and the randomness generation speed is 5.10×10^4 bits per second. In the second model, by utilizing single-photon antibunching effect, only single-photon events are used to extract randomness, and the randomness output speed is 4.74×10^4 bits per second, which is in the same magnitude as the first model. Moreover, in the second model, the single-photon QRNG can be considered a source-independent random number generator, which does not require the trust of light sources. In the third method, only tuple detection events below the unity line are considered raw bits. The security level is highest in the third model, but the randomness output speed drops to 34.37 bits per second. Taking both speed and security into account, the second model is an ideal choice for a single-photon QRNG.

In Chapter 5, a QRNG model based on a delayed-choice experiment to get quantum random numbers without the fair sampling assumption is constructed. We use the wave-particle duality in a delayed-choice experiment to guarantee that photons arrive at the detectors in superposition states, so the fair sampling assumption is no longer needed. By applying our model to the delayed-choice experiment performed by Jacques *et al.* [1], $1,124$ uniformally distributed random bits can be obtained per second.

In Chapter 6, we show how to certify quantum randomness from a loophole-free Bell test data [62] by Bell's theorem [2] and RSP-dimension witness [83, 84, 4, 3]. With the CHSH inequality [73] in Bell's theorem, the min-entropy of quantum randomness can be quantified in a DI-way [82]. Using the RSP-dimension witness, SDI random numbers can be extracted. Compared with the DI model in [82], our SDI model increases the randomness from 0.956×10^{-2} bits per event to 0.153 bits per event, and correspondingly raise randomness output speed from 2.54 bits per day to 40.63 bits per day, which is one important step towards the practical use of the Bell test in randomness generation.

In Chapter 7, a QRNG based on a nuclear spin system inside an NV center is studied, including two QRNG models. The first model directly applies the two-dimensional dimension witness protocol in [4]. In our second dimension witness model, we develop a protocol based on a three-dimensional dimension witness W_3. Our W_3 based QRNG protocol can be applied to any P&M scenario with six input states, three measurement bases, and binary results. Also, from our experimental data, the model based on W_3 can bound 53% more random bits than the W_2 model. The randomness output speed is 1.33 bits per second in the W_3 model, which is higher than 0.87 bits per second of the W_2 model that from [4].

8.2 Outlook

Quantum random number generation has evolved into a relatively mature quantum technology. Optical QRNGs, in particular, have achieved generation rates on the order of gigabits per second for over a decade [185, 186, 69], and the pursuit of higher generation rates continues [187, 188, 189]. Concurrently, some optical QRNGs have transitioned from laboratory settings to commercial products (ID Quantique, QRBG121, Quintessence Labs). However, optical QRNGs can fail due to device malfunction or external attacks. Therefore, developing a reliable and fast QRNG capable of self-testing is not only of academic interest but also has practical implications. As the demand for cryptocurrencies and quantum key distribution grows, the need for true and reliable randomness will also increase.

The four QRNGs investigated in this thesis highlight various approaches to constructing reliable QRNGs by exploiting distinct quantum phenomena. Each QRNG can be enhanced through future developments: As single-photon sources become more prevalent, sources with higher emission rates and improved quality will be developed. Additionally, the detection efficiency of single-photon detectors will improve. The single-photon QRNG will benefit from these advancements, likely eliminating the need for the fair-sampling assumption. In the delayed-choice QRNG, the experiment is conducted with single-photons, which are challenging to interfere with in space-like separation using long fibers. Consequently, the interference pattern is unstable; however, as single-photon sources and optical fibers advance, maintaining the interference patterns of single-photons will become more manageable, bolstering the robustness of the delayed-choice QRNG scheme. For the Bell test QRNG, the current technological requirements are stringent; as quantum technologies progress, these requirements will become more achievable, bringing the DI-QRNG closer to practical use.

The nuclear spin-based QRNG benefits from high state preparation and measurement fidelities. However, these high fidelities are achieved through sophisticated experimental devices and complex operations. To make this QRNG more practical, future developments must focus on creating more portable setups and simplifying the operation process.

In summary, research on reliable QRNGs will persist, with advancements occurring in two primary ways. Firstly, the development of mathematical protocols used to quantify the entropy of generated raw random bits will advance. With more refined theoretical protocols, greater quantum randomness will be quantified from the same experimental data compared to existing protocols. Secondly, the evolution of quantum technologies and platforms will contribute to generating quantum randomness more robustly and rapidly. The ongoing progress in quantum random number generation will lead to more secure and efficient applications across various fields, further solidifying its importance in both academia and industry.

Appendix

We add appendices here for each chapter.

Appendix for Chapter 4

Detection probability and conditional probability.

For the convenience of description, the detector in the transmitted arm is named detector A (events assigned to raw bit “0”), and the detector in the reflected arm is named detector B (events assigned to raw bit “1”). Next, we deduce all the conditional probabilities from the experimental parameters. This does not describe the true probabilities but reflects the frequencies of the occurring singles and tuples.

Take $p(A|A)$ as an example. $p(A|A)$ means the probability of a subsequent photon to be detected in detector A when detector A has already detected a previous photon event. Let η_{A} be the detection efficiency of detector A, $\tau_{\mathrm{dead}}^{\mathrm{A}}$ be the dead-time of detector A, and T be the transmission coefficient of the beam-splitter. When detector A clicks, it is in its dead time. $\int_0^{\tau_{\mathrm{dead}}^{\mathrm{A}}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau$ means the probability that the next incident photon is in the dead-time of detector A, then the probability of this incident photon outside its dead-time is $1-\int_0^{\tau_{\mathrm{dead}}^{\mathrm{A}}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau$. When the incident photon is outside the dead-time of detector A, it has probability T to be transmitted to detector A, and detector A has probability η_{A} to detect this photon, so $p(A|A)$ could be written as

$$p(A|A) = \big(1-\int_0^{\tau_{\mathrm{dead}}^{\mathrm{A}}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau\big)\eta_{\mathrm{A}}T\ . \tag{8.1}$$

Similarly, the parametric equation for $p(B|A)$ is

$$p(B|A) = \eta_{\mathrm{B}}R\big(1 - \underbrace{\eta_{\mathrm{B}}R \int_0^{\frac{\tau_{\mathrm{dead}}^{\mathrm{B}}}{2}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau \int_0^{\frac{\tau_{\mathrm{dead}}^{\mathrm{B}}}{2}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau\big)}_{\text{probability that detector B is in its dead-time}} . \tag{8.2}$$

where η_{B} is the detection efficiency of detector B, $\tau_{\mathrm{dead}}^{\mathrm{B}}$ is the dead-time of detector B, and R is the reflection coefficient. This gives us the equation of $p(B|A)$, which means that when detector A detects a photon event, the probability of detector B detecting a subsequent photon event.

The formula in the underbrace means the probability of detector B is not in its dead-time when a photon shoots into the beamsplitter. This probability is an estimation, which is based on the assumption $\tau_{\mathrm{dead}}^{\mathrm{A}} \approx \tau_{\mathrm{dead}}^{\mathrm{B}}$. Before detector A clicks, the previous photon event may be on detector A or B. If it is on detector A, it will not affect the conditional probability $p(B|A)$ since when detector A clicks two times, detector B is ready to detect a photon event; if the previous photon event is on detector B, then after detector A's click, detector B still has a probability to be in its dead-time when the next photon comes into the beamsplitter. Inside the brace, the half of the dead-time of detector B is a simplified version of the above probability analysis, where $\eta_{\mathrm{B}}R\int_0^{\frac{\tau_{\mathrm{dead}}^{\mathrm{B}}}{2}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau$ gives us the probability that the previous photon event fires on detector B within the half of the dead-time of detector B, and $\int_0^{\frac{\tau_{\mathrm{dead}}^{\mathrm{B}}}{2}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau$ is the probability that the next incident photon is inside the half of the dead-time of detector B.

$p(B|B)$ and $p(A|B)$ can be derived analogously.

The equation of $p(A)$ is $p(A) = r_{\mathrm{A}}/r_{\mathrm{total}}$, r_{A} is the click rate of detector A. r_{A} is defined as following

$$\begin{aligned} r_{\mathrm{A}} =& \eta_{\mathrm{A}}TI_{\mathrm{in}} - \underbrace{\frac{(\eta_{\mathrm{A}}TI_{\mathrm{in}})}{2} \times \frac{(\eta_{\mathrm{A}}TI_{\mathrm{in}})}{2} \int_0^{\tau_{\mathrm{dead}}^{\mathrm{A}}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau}_{\text{rate of two clicks within the dead-time of detector A}} \\ =& \eta_{\mathrm{A}}TI_{\mathrm{in}} - \frac{(\eta_{\mathrm{A}}TI_{\mathrm{in}})^2 \int_0^{\tau_{\mathrm{dead}}^{\mathrm{A}}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau}{4} , \end{aligned} \tag{8.3}$$

where I_{in} is the rate of the incident photon, and $\eta_{\mathrm{A}}TI_{\mathrm{in}}$ means the click rate when detector A would have no dead-time. The latter part of the equation is the probability of two subsequent events in detector A having a time distance that is smaller than the dead time of detector A. For detector B, a similar equation of r_{B} is derived, and we get

$$\begin{aligned}
p(A) &= \frac{r_{\mathrm{A}}}{r_{\mathrm{A}}+r_{\mathrm{B}}} \\
&= \frac{\eta_{\mathrm{A}}TI_{\mathrm{in}} - \frac{(\eta_{\mathrm{A}}TI_{\mathrm{in}})^2\int_0^{\tau_{\mathrm{dead}}^{\mathrm{A}}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau}{4}}{\eta_{\mathrm{A}}TI_{\mathrm{in}} - \frac{(\eta_{\mathrm{A}}TI_{\mathrm{in}})^2\int_0^{\tau_{\mathrm{dead}}^{\mathrm{A}}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau}{4} + \eta_{\mathrm{B}}RI_{\mathrm{in}} - \frac{(\eta_{\mathrm{B}}RI_{\mathrm{in}})^2\int_0^{\tau_{\mathrm{dead}}^{\mathrm{B}}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau}{4}} \\
&= \frac{\eta_{\mathrm{A}}T - \frac{(\eta_{\mathrm{A}}T)^2 I_{\mathrm{in}}\int_0^{\tau_{\mathrm{dead}}^{\mathrm{A}}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau}{4}}{\eta_{\mathrm{A}}T - \frac{(\eta_{\mathrm{A}}T)^2 I_{\mathrm{in}}\int_0^{\tau_{\mathrm{dead}}^{\mathrm{A}}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau}{4} + \eta_{\mathrm{B}}R - \frac{(\eta_{\mathrm{B}}R)^2 I_{\mathrm{in}}\int_0^{\tau_{\mathrm{dead}}^{\mathrm{B}}} g_{\mathrm{fit}}^{(2)}(\tau)d\tau}{4}} \, .
\end{aligned} \tag{8.4}$$

With all the above equations, the parametric expression of $H_\infty(X|Y)$ could be deduced.

The error bound of the conditional min-entropy

In this subsection, the error bound of the conditional min-entropy is given. Let us mention some properties of the error bound. Since $p(A)+p(B)=1$ is always fulfilled, we have $\Delta_{p(A)} = -\Delta_{p(B)}$. Also, for conditional probabilities $p(A|A), p(B|A), p(A|B), p(B|B)$, $p(A|A)+p(B|A)=1$ and $p(A|B)+p(B|B)=1$, this means $\Delta_{p(A|A)} = -\Delta_{p(B|A)}$ and $\Delta_{p(A|B)} = -\Delta_{p(B|B)}$. Since $p(AB) = p(BA)(p(BA) = p(AB)$ is satisfied under the condition that the experimental devices does not change over time), and $\Delta_{p(A)} = -\Delta_{p(B)}$, it is easy to derive the relationship $\Delta_{p(A|A)} = -\Delta_{p(B|A)} = \Delta_{p(A|B)} = -\Delta_{p(B|B)}$.

The conditional min-entropy in our case is defined in 4.5. Since $p(BA) = p(AB)$,

$$H_\infty(X|Y) = -\log_2\big(\max\{p(A)-p(AB), p(AB)\} + \max\{p(AB), 1-p(A)-p(AB)\}\big) \, . \tag{8.5}$$

There are four different conditions for $H_\infty(X|Y)$

$$H_\infty(X|Y) = \begin{cases} -\log_2(p(A)) & p(A) - p(AB) \geq p(AB) \quad \mathit{and} \quad p(AB) \geq 1 - p(A) - p(AB)\,, \\ -\log_2(p(B)) & p(A) - p(AB) \leq p(AB) \quad \mathit{and} \quad p(AB) \leq 1 - p(A) - p(AB)\,, \\ -\log_2(2p(AB)) & p(A) - p(AB) \leq p(AB) \quad \mathit{and} \quad p(AB) \geq 1 - p(A) - p(AB)\,, \\ -\log_2(1-2p(AB)) & p(A) - p(AB) \geq p(AB) \quad \mathit{and} \quad p(AB) \leq 1 - p(A) - p(AB)\,. \end{cases} \tag{8.6}$$

No matter which condition $H_\infty(X|Y)$ is in, there is only one variable in it. Then a more conservative conditional min-entropy could be written as

$$H_\infty(X|Y) = -\log_2(f(p) + \Delta_{f(p)}) \tag{8.7}$$

where $f(p) = \max\{p(A), p(B), 2p(AB), 1-2p(AB)\}$.

The equation of $p(A)$ is Eqn. 8.4, $p(A)$ is affected by the transmission coefficient T, the rate of incident photon I_{in}, the detection efficiency η_{A}, η_{B}, and the dead-time τ^A_{dead}, τ^B_{dead} of the two detectors. The error bound of each parameter is, δ_T, $\delta_{\eta_{\text{A}}}$,$\delta_{\eta_{\text{B}}}$, $\delta_{\tau^A_{\text{dead}}}$,$\delta_{\tau^B_{\text{dead}}}$, and $\delta_{I_{\text{in}}}$. According to the error propagation, the error bound of $p(A)$ is

$$\begin{aligned} \Delta_{p(A)} = \Bigg(& (\frac{\partial p(A)}{\partial T}\delta_T)^2 + (\frac{\partial p(A)}{\partial \eta_{\text{A}}}\delta_{\eta_{\text{A}}})^2 + (\frac{\partial p(A)}{\partial \eta_{\text{B}}}\delta_{\eta_{\text{B}}})^2 + \\ & (\frac{\partial p(A)}{\partial \tau^A_{\text{dead}}}\delta_{\tau^A_{\text{dead}}})^2 + (\frac{\partial p(A)}{\partial \tau^B_{\text{dead}}}\delta_{\tau^B_{\text{dead}}})^2 + (\frac{\partial p(A)}{\partial I_{\text{in}}}\delta_{I_{\text{in}}})^2 \Bigg)^{\frac{1}{2}}\,. \end{aligned} \tag{8.8}$$

The equation of $p(AB)$ is $p(AB) = p(A)p(B|A)$. From 8.2 and 8.4, we know that multiple parameters affect $p(AB)$, including τ^A_{dead}, τ^B_{dead}, η_{A}, η_{B}, T, and I_{in}, similarly, $\Delta_{p(AB)}$ is

$$\begin{aligned} \Delta_{p(AB)} = \Bigg(& (\frac{\partial p(AB)}{\partial T}\delta_T)^2 + (\frac{\partial p(AB)}{\partial \eta_{\text{A}}}\delta_{\eta_{\text{A}}})^2 + (\frac{\partial p(AB)}{\partial \eta_{\text{B}}}\delta_{\eta_{\text{B}}})^2 + \\ & (\frac{\partial p(AB)}{\partial \tau^A_{\text{dead}}}\delta_{\tau^A_{\text{dead}}})^2 + (\frac{\partial p(AB)}{\partial \tau^B_{\text{dead}}}\delta_{\tau^B_{\text{dead}}})^2 + (\frac{\partial p(AB)}{\partial I_{\text{in}}}\delta_{I_{\text{in}}})^2 \Bigg)^{\frac{1}{2}}\,. \end{aligned} \tag{8.9}$$

Then from Eqn. 8.7, 8.8, 8.9, the conservative $H_\infty(X|Y)$ could be calculated.

For the second model, there is one more parameter p_e, which represents the probability of detecting an uncorrelated background noise event. The conditional min-entropy for the second model is

$$\begin{aligned} H_\infty(X|Y) &= -\log_2\Big(p_e + (1-p_e)\big(\sum_y p(y)\max_x\{p(x|y)\}\big)\Big) \\ &= -\log_2\Big(p_e + (1-p_e)f(p)\Big) . \end{aligned} \tag{8.10}$$

where $f(p) = \max\{p(A), p(B), 2p(AB), 1-2p(AB)\}$. A more conservative $H_\infty(X|Y)$ for this model is

$$H_\infty(X|Y) = -\log_2\big(p_{eq} + \Delta_{p_{eq}}\big) \tag{8.11}$$

where $p_{eq} = p_e + (1-p_e)f(p)$, then $\Delta_{p_{eq}} = \sqrt{(\frac{\partial p_{eq}}{\partial p_e}\Delta_{p_e})^2 + (\frac{\partial p_{eq}}{\partial f(p)}\Delta_{f(p)})^2}$, where $\Delta_{f(p)}$ is derived from 8.8 and 8.9, and Δ_{p_e} is shown in 8.12.

Note that the uncertainty of $g^{(2)}_{\text{fit}}(0)$ will affect the fraction of background noise, thus affect p_e. Since $p_e = 1 - s = 1 - \sqrt{1 - g^{(2)}_{\text{fit}}(0)}$, according to the propagation of uncertainty, the uncertainty of p_e is

$$\Delta_{p_e} = \frac{1}{2\sqrt{1-g^{(2)}_{\text{fit}}(0)}}\Delta_{g^{(2)}_{\text{fit}}(0)} . \tag{8.12}$$

where $\Delta_{g^{(2)}_{\text{fit}}(0)}$ is derived in the following subsection(see Eqn. 8.22).

The error bound of the classical limit line

In the third model, the extractable quantum randomness in the raw data is determined by the single photon start-stop event count rate under the classical limit line [159, 161, 109]

$$
\begin{aligned}
r_{\mathrm{rand}} &= r_{\mathrm{A}}\sqrt{1-g^{(2)}_{\mathrm{fit}}(0)} \times r_{\mathrm{B}}\sqrt{1-g^{(2)}_{\mathrm{fit}}(0)} \times \int_{-t}^{t} g^{(2)}_{\mathrm{fit}}(\tau)d\tau \\
&= (1-g^{(2)}_{\mathrm{fit}}(0)) \times (\eta_{\mathrm{A}}TI_{\mathrm{in}} - \frac{(\eta_{\mathrm{A}}TI_{\mathrm{in}})^2 \int_0^{\tau^{\mathrm{A}}_{\mathrm{dead}}} g^{(2)}_{\mathrm{fit}}(\tau)d\tau}{4})(\eta_{\mathrm{B}}RI_{\mathrm{in}} - \frac{(\eta_{\mathrm{B}}RI_{\mathrm{in}})^2 \int_0^{\tau^{\mathrm{B}}_{\mathrm{dead}}} g^{(2)}_{\mathrm{fit}}(\tau)d\tau}{4}) \\
&\quad \times \int_{-t}^{t} g^{(2)}_{\mathrm{fit}}(\tau)d\tau \\
&= (1-g^{(2)}_{\mathrm{fit}}(0)) \times (\eta_{\mathrm{A}}T - \frac{(\eta_{\mathrm{A}}T)^2 I_{\mathrm{in}} \int_0^{\tau^{\mathrm{A}}_{\mathrm{dead}}} g^{(2)}_{\mathrm{fit}}(\tau)d\tau}{4})(\eta_{\mathrm{B}}R - \frac{(\eta_{\mathrm{B}}R)^2 I_{\mathrm{in}} \int_0^{\tau^{\mathrm{B}}_{\mathrm{dead}}} g^{(2)}_{\mathrm{fit}}(\tau)d\tau}{4}) \\
&\quad \times I_{\mathrm{in}}^2 \int_{-t}^{t} g^{(2)}_{\mathrm{fit}}(\tau)d\tau \ .
\end{aligned}
\tag{8.13}
$$

where t satisfies $g^{(2)}_{\mathrm{fit}}(t) = 1$. And the quantum fraction of the raw bits is defined as $r_{\mathrm{rand}}/r_{\mathrm{total}}$, then p_c, the classical noise probability, is $p_c = 1 - r_{\mathrm{rand}}/r_{\mathrm{total}}$. Under the classical limit line, the events 0-1 is taken as random bit **0**, and 1-0 is taken as **1**, the conditional min-entropy in this case is defined as

$$
\begin{aligned}
H_\infty(\mathbf{X}|\mathbf{Y}) &= -\log_2\Big(p_c + (1-p_c)\big(\max\{p(\mathbf{00}), p(\mathbf{01})\} + \max\{p(\mathbf{10}), p(\mathbf{11})\}\big)\Big) \\
&= -\log_2\Big(p_c + (1-p_c)f(\mathbf{p})\Big) \ .
\end{aligned}
\tag{8.14}
$$

where $f(\mathbf{p}) = \max\{p(\mathbf{0}), p(\mathbf{1}), 2p(\mathbf{01}), 1-2p(\mathbf{01})\}$.

For the convenience of description, without losing generality, we associate event pair “AB" to random bit **0**, and “BA" to **1**. For probabilities $p(\mathbf{0})$ and $p(\mathbf{1})$, there are two different situations. The first situation: when t is larger than the half of the dead-time of the detectors, we have

$$p(\mathbf{0}) = p(AB) = p(A)\eta_{\mathrm{B}}R\Big(\int_0^t g^{(2)}_{\mathrm{fit}}(\tau)d\tau - \eta_{\mathrm{B}}R \int_0^{\frac{\tau^{\mathrm{B}}_{\mathrm{dead}}}{2}} g^{(2)}_{\mathrm{fit}}(\tau)d\tau \times \underbrace{\int_0^{\frac{\tau^{\mathrm{B}}_{\mathrm{dead}}}{2}} g^{(2)}_{\mathrm{fit}}(\tau)d\tau}_{\text{incident photon within } \tau^{\mathrm{B}}_{\mathrm{dead}}/2.} \Big) . \tag{8.15}$$

The probability here is very similar to 8.2, except we only consider short-time related photon events in this situation, so we replace '1' in 8.2 with $\int_0^t g^{(2)}_{\mathrm{fit}}(\tau)d\tau$ here, where $p(A)$ is in 8.4. The other situation is when t is smaller than the half of the dead-time of each detector, we need to change the formula inside the underbrace to $\int_0^t g^{(2)}_{\mathrm{fit}}(\tau)d\tau$, then

$$p(\mathbf{0}) = p(AB) = p(A)\eta_{\mathrm{B}}R \int_0^t g^{(2)}_{\mathrm{fit}}(\tau)d\tau \big(1 - \eta_{\mathrm{B}}R \int_0^{\frac{\tau^{\mathrm{B}}_{\mathrm{dead}}}{2}} g^{(2)}_{\mathrm{fit}}(\tau)d\tau\big) . \tag{8.16}$$

The equation of $p(\mathbf{1})$ can be deduced in a similar way.

For the photon events under the classical limit line, the events pair **0** or **1** is are much less correlated than previous models, they can be treated as independent events, so we have

$$\begin{aligned} p(\mathbf{00}) &= p(\mathbf{0})p(\mathbf{0}) , \\ p(\mathbf{01}) &= p(\mathbf{0})p(\mathbf{1}) , \\ p(\mathbf{10}) &= p(\mathbf{1})p(\mathbf{0}) , \\ p(\mathbf{11}) &= p(\mathbf{1})p(\mathbf{1}) . \end{aligned} \tag{8.17}$$

Next we calculate the conservative conditional min-entropy in this model, similar to the second model, we have

$$H_\infty(\mathbf{X}|\mathbf{Y}) = -\log_2\big(p_{cq} + \Delta_{p_{cq}}\big) . \tag{8.18}$$

where $p_{cq} = p_c + (1 - p_c)f(\mathbf{p})$, then $\Delta_{p_{cq}} = \sqrt{(\frac{\partial p_{cq}}{\partial p_c}\Delta_{p_c})^2 + (\frac{\partial p_{cq}}{\partial f(\mathbf{p})}\Delta_{f(\mathbf{p})})^2}$. From $p_c = 1 - r_{\mathrm{rand}}/r_{\mathrm{total}}$, we get

$$p_c = 1 - \frac{(1 - g^{(2)}_{\mathrm{fit}}(0)) \times r_{\mathrm{A}} \times r_{\mathrm{B}} \times \int_{-t}^t g^{(2)}_{\mathrm{fit}}(\tau)d\tau}{r_{\mathrm{A}} + r_{\mathrm{B}}} . \tag{8.19}$$

and the equation for $f(\mathbf{p})$ is in 8.16, 8.15 and 8.17. From the equations of p_c and $f(\mathbf{p})$, we can see that they are dependent on some same parameters, including the dead-time of the two detectors, the detection efficiencies, and the beam-splitter ratio etc. This means that they are not independent from each other, so $\Delta_{p_{cq}} \neq \sqrt{(\frac{\partial p_{cq}}{\partial p_c}\Delta_{p_c})^2 + (\frac{\partial p_{cq}}{\partial f(\mathbf{p})}\Delta_{f(\mathbf{p})})^2}$, $\Delta_{p_{cq}}$ should be derived directly from the experimental parameters

$$\begin{aligned}\Delta_{p_{cq}} = \Bigg(&(\frac{\partial p_c}{\partial T}\delta_T)^2 + (\frac{\partial p_c}{\partial \eta_\mathrm{A}}\delta_{\eta_\mathrm{A}})^2 + (\frac{\partial p_c}{\partial \eta_\mathrm{B}}\delta_{\eta_\mathrm{B}})^2 + (\frac{\partial p_c}{\partial \tau^A_\mathrm{dead}}\delta_{\tau^A_\mathrm{dead}})^2 + \\ &(\frac{\partial p_c}{\partial \tau^B_\mathrm{dead}}\delta_{\tau^B_\mathrm{dead}})^2 + (\frac{\partial p_c}{\partial I_\mathrm{in}}\delta_{I_\mathrm{in}})^2 + \frac{\partial p_c}{\partial t}\delta_t)^2\Bigg)^{\frac{1}{2}} .\end{aligned} \tag{8.20}$$

where t satisfies $g^{(2)}_\mathrm{fit}(t) = 1$. Next we derive δ_t. In our case, δ_t is characterized by the classical limit line. The classical limit line is determined by the normalization factor of the experimental anti-bunching curve. The normalization factor N_norm is calculated by

$$N_\mathrm{norm} = r_\mathrm{A} r_\mathrm{B} \tau_\mathrm{rs} T_\mathrm{total}$$

where τ_rs is the timing resolution of the start-stop event, T_total is the total integration time (the running time of the experiment). N_norm can be determined by multiple parameters, including the detection efficiency and dead-time of each detector, and the reflection and transmission coefficients. According to the propagation of uncertainty, we get the uncertainty of N_norm

$$\begin{aligned}\Delta_\mathrm{norm} = \Bigg(&(\frac{\partial N_\mathrm{norm}}{\partial R}\delta_R)^2 + (\frac{\partial N_\mathrm{norm}}{\partial \eta_\mathrm{A}}\delta_{\eta_\mathrm{A}})^2 + \\ &(\frac{\partial N_\mathrm{norm}}{\partial \eta_\mathrm{B}}\delta_{\eta_\mathrm{B}})^2 + (\frac{\partial N_\mathrm{norm}}{\partial I_\mathrm{in}}\delta_{I_\mathrm{in}})^2 + \\ &(\frac{\partial N_\mathrm{norm}}{\partial \tau^\mathrm{A}_\mathrm{dead}}\delta_{\tau^\mathrm{A}_\mathrm{dead}})^2 + (\frac{\partial N_\mathrm{norm}}{\partial \tau^\mathrm{B}_\mathrm{dead}}\delta_{\tau^\mathrm{B}_\mathrm{dead}})^2\Bigg)^{\frac{1}{2}} .\end{aligned}$$

Then the uncertainty of the classical limit line (i.e. $g^{(2)}_\mathrm{fit}(\tau) = 1$) amounts to

$$\Delta_1 = 1 \times \frac{\Delta_\mathrm{norm}}{N_\mathrm{norm}} . \tag{8.21}$$

and the uncertainty of the background line $g^{(2)}_\mathrm{fit}(0)$ is

$$\Delta_{g_{\mathrm{fit}}^{(2)}(0)} = g_{\mathrm{fit}}^{(2)}(0)\frac{\Delta_{\mathrm{norm}}}{N_{\mathrm{norm}}} \,. \tag{8.22}$$

From the uncertainty of classical limit line, $\delta_t = t - t'$ can be deduced, where t' satisfies the equation $g_{\mathrm{fit}}^{(2)}(t') = 1 - \Delta_1$. Then $\Delta_{p_{cq}}$ can be derived.

Appendix for Chapter 5

Photon statistics at polarized beamsplitters

Here we give an analytical equation to show how photon interference happens in the Mach-Zehnder interferometer with polarizing beamsplitters (PBSs), as shown in our experimental scheme in Fig. 5.1. For the convenience of illustration, part of our scheme is reshown here.

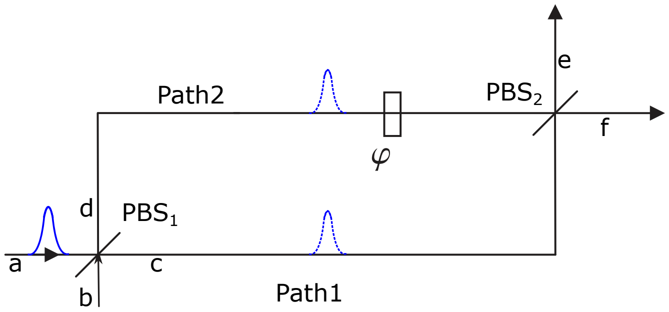


Figure 8.1: MZI with polarizing beamsplitters. In the delayed-choice experiment by Jacques *et al.*, they use polarizing beamsplitters PBS$_1$ and PBS$_2$. Path 1 and Path 2 are fibers with the same length.

Polarized beamsplitters can reflect only vertical polarized photons and transmit horizontally polarized photons. The polarization adds an additional degree of freedom to the photon. Suppose photon incidents from the port (a), and then after PBS$_1$, the photon is in state

$$|a\rangle \xrightarrow{\mathrm{PBS}_1} \frac{1}{\sqrt{2}}(|c_H\rangle + i\,|d_V\rangle) \tag{8.23}$$

where H, V represent horizontal and vertical polarization directions. Then photon travels in path 1 and path 2 and recombines in the second PBS. State $|c\rangle$ and $|d\rangle$ evolves in the following way

$$\begin{aligned} |c\rangle &\xrightarrow{\text{PBS}_2} i\,|f_H\rangle \\ |d\rangle &\xrightarrow{\varphi} e^{i\varphi}\,|d\rangle \xrightarrow{\text{PBS}_2} ie^{i\varphi}\,|e_V\rangle \end{aligned} \tag{8.24}$$

Then the total process of state $|a\rangle$ is

$$|a\rangle \xrightarrow{\text{PBS}_1} \frac{1}{\sqrt{2}}(|c_H\rangle + i\,|d_V\rangle) \xrightarrow{\varphi} \frac{1}{\sqrt{2}}(|c_H\rangle + ie^{i\varphi}\,|d_V\rangle \xrightarrow{\text{PBS}_2} \frac{1}{\sqrt{2}}(|f_H\rangle + ie^{i\varphi}\,|e_V\rangle) \tag{8.25}$$

Apparently, after PBS$_2$, there is no interference since photons in the two paths are distinguishable. Now an EOM is put after ports (e) and (f), and voltage V_π is applied. The EOM is behaving as a half-wave plate with a fast axis at angle $22.5°$, the state $|f_H\rangle$ and $|e_V\rangle$ will be changed to

$$\begin{aligned} |f_H\rangle &\xrightarrow{\text{EOM}} \frac{1}{\sqrt{2}}(|f_H\rangle + |f_V\rangle) \\ |e_V\rangle &\xrightarrow{\text{EOM}} \frac{1}{\sqrt{2}}(|e_H\rangle - |e_V\rangle) \end{aligned} \tag{8.26}$$

As shown in Fig. 8.2, after the EOM, beam (e) and (f) are recombined in PBS$_3$. Then we have

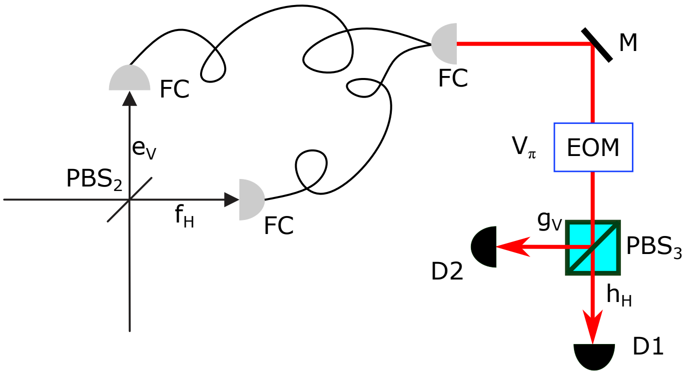


Figure 8.2: Beams recombine after PBS$_2$.

$$\begin{aligned} |a\rangle &\xrightarrow{\text{PBS}_1} \frac{1}{\sqrt{2}}(|c_H\rangle + i\,|d_V\rangle) \xrightarrow{\varphi} \frac{1}{\sqrt{2}}(|c_H\rangle + ie^{i\varphi}\,|d_V\rangle \xrightarrow{\text{PBS}_2} \frac{1}{\sqrt{2}}(|f_H\rangle + ie^{i\varphi}\,|e_V\rangle) \\ &\xrightarrow{EOM} \frac{1}{2}\left(f_H + f_V + ie^{i\varphi}e_H - ie^{i\varphi}e_V\right) \xrightarrow{\text{PBS}_3} \frac{1}{2}(1 + ie^{i\varphi})h_H + \frac{1}{2}(1 - ie^{i\varphi})g_V \end{aligned} \tag{8.27}$$

The probability of finding a photon in beam (g) and (h) is

$$\begin{aligned} p_g &= \frac{1}{2}(1+\sin\varphi) \\ p_h &= \frac{1}{2}(1-\sin\varphi) \end{aligned} \tag{8.28}$$

With the change of phase-shift φ, the interference patterns can be observed in D_1 and D_2.

Appendix for Chapter 6

Guessing probability of $\mathrm{p(ab|xy)}$ under SDI conditions.

With the assumption being settled in Chapter 6, the guessing probability can now be derived, and then with the guessing probability, the min-entropy in the raw bits can be quantified.

Under SDI conditions, according to the assumptions, Alice's and Bob's measurement settings are independent of each other, and the internal states of the experimental devices are not affected by previous results during the experimental run. Due to the requirement that subsequent measurements are i.i.d., their choices of measurements are uniformly random; thus, each combination of x and y occurs with probability 1/4. Then the guessing probability $p_g(ab|xy)$ can be defined as

$$p_g(ab|xy) = \frac{1}{4}\sum_{x,y}\max_{a,b} p(ab|xy) \tag{8.29}$$

The right part of this equation satisfies

$$\begin{aligned}
&\frac{1}{4}\sum_{x,y}\max_{a,b} p(ab|xy)\\
&=\frac{1}{4}\sum_{x,y}\max_{a,b} p(a|x)p(b|x,a,y)\\
&\leq\frac{1}{4}\sum_{x,y}\max_{a,b} p(a|x)\max_{a,b} p(b|x,a,y)\\
&\leq\frac{1}{2}\sum_{x}\max_{a} p(a|x)\frac{1}{2}\sum_{y}\max_{x,a,b} p(b|(x,a),y)\\
&\leq\max_{x,a} p(a|x)\frac{1}{2}\sum_{y}\max_{x,a,b} p(b|(x,a),y),
\end{aligned} \tag{8.30}$$

where the upper bound of $\frac{1}{2}\sum_y \max_{x,a,b} p(b|(x,a),y)$ can be obtained from [4], and it is

$$\frac{1}{2}\sum_{y}\max_{x,a,b} p(b|(x,a),y) \leq \frac{1}{2}\left(1+\sqrt{\frac{1+\sqrt{1-W_B^2}}{2}}\right). \tag{8.31}$$

The next step is to find an upper bound of $p(a|x)$ with the given W_B.

In the ideal case, with a perfect maximally entangled state and a perfect measurement procedure at Alice's side, the probability distribution for a local measurement outcome a is $p(a|x) = 0.5$. However, in reality, the experimental devices will have imperfections, and thus, the probabilities might deviate from the perfect value of 0.5, and this deviation can be connected to W_B.

Each combination of a and x is associated with a state $\rho_{a|x}$ (see, Eqn (6.3)). Since these are qubit states, we write them as

$$\rho_{a|x} = \frac{\mathbf{I}_2 + \vec{s}_{a|x}\cdot\vec{\sigma}}{2}, \tag{8.32}$$

where $\vec{s}_{a|x}$ is a Bloch vector, $\vec{\sigma} = \sigma_x\hat{\imath} + \sigma_y\hat{\jmath} + \sigma_z\hat{k}$ is the Pauli vector, and $\mathbf{I}_2$ is a 2-dimension unity matrix. Additionally, the measurement operator for Bob's side is

$$M^B_{b|y} = \frac{c_y\mathbf{I}_2 + \vec{T}_y\cdot\vec{\sigma}}{2}, \tag{8.33}$$

where $|c_y| \leq 1$, and $\vec{T}_y$ is a Bloch vector. Inserting the state $\rho_{a|x}$ (Eqn. (8.32)) and the measurement operator $M^B_{b|y}$ (Eqn. (8.33)) in Eqn.(2) allows to calculate the probability $p(b|a,x,y) = p(b|x',y)$ for a measurement outcome b depending on a, x, and y. In the following we only consider $b = 1$ and formulate this depending on a, x, and y. Define $p(b = 1|(x,a),y) = p((a|x),y)$, and then the matrix elements of the dimension witness (Eqn.(3)) take the form

$$p((a = 0|x), y) - p((a = 1|x), y) = \mathrm{Tr}[(\rho_{a=0|x} - \rho_{1|x})M^B_{b=1|y}] = \vec{S}_x \cdot \vec{T}_y, \tag{8.34}$$

where $\vec{S}_x = (\vec{s}_{a=0|x} - \vec{s}_{a=1|x})/2$.

To get the upper bound of $p(a|x)$ with the given W_B, suppose the entangled state shared between Alice and Bob is $|\Psi^+_\theta\rangle = \cos\theta\,|01\rangle - \sin\theta\,|10\rangle$. For this state, when Alice performs her two measurements $\hat{x}$ and $\hat{z}$ on this state, we have $\max p(a|x) \leq \cos^2\theta$. Correspondingly, Bob will get four states on his side. Then according to Eqn. (8.34), and Eqn.(8) in [83], we have

$$\begin{aligned} W_B = &\begin{vmatrix} \vec{S}_0 \cdot \vec{T}_0 & \vec{S}_1 \cdot \vec{T}_0 \\ \vec{S}_0 \cdot \vec{T}_1 & \vec{S}_1 \cdot \vec{T}_1 \end{vmatrix} \\ = &\left(\vec{S}_0 \times \vec{S}_1\right) \cdot \left(\vec{T}_0 \times \vec{T}_1\right) \\ \leq &\left|\vec{S}_0 \times \vec{S}_1\right| \cdot \left|\vec{T}_0 \times \vec{T}_1\right| \leq 1. \end{aligned} \tag{8.35}$$

where $\left|\vec{S}_0 \times \vec{S}_1\right| = \sin 2\theta$. Suppose the measurement settings of Bob's side are in ideal conditions, which means $\left|\vec{T}_0 \times \vec{T}_1\right| = 1$. Then $W_B \leq \sin 2\theta$. Considering $\max p(a|x) \leq \cos^2\theta$, the upper bound of $p(a|x)$ can be quantified as

$$\max_{a,x} p(a|x) \leq \frac{1 + \sqrt{1 - W_B^2}}{2} \tag{8.36}$$

The above derivation process leads to the maximum $p(a|x)$ with the given W_B value. Of course, the deviation of W_B from the optimal value 1 might have other causes other than the non-maximal pure entangled state. However, compared to our analysis above, none of them can produce a larger upper bound of $p(a|x)$. Two more different scenarios may lead to a larger guessing probability with given W_B. The first scenario is that we have a mixed entangled state, which is:

$$\rho_z = \underbrace{\frac{1-z}{4}\mathbf{I}}_{Noise} + \underbrace{z\,|\Psi_\theta^+\rangle\,\langle\Psi_\theta^+|}_{|\Psi_\theta^+\rangle} \tag{8.37}$$

where $z \in (0,1)$ and $|\Psi_\theta^+\rangle = \cos\theta\,|01\rangle - \sin\theta\,|10\rangle$. Suppose the dimension witness of this system is still W_B, and let the dimension witness of the $|\Psi_\theta^+\rangle$ part be W_B'. Then according to [83], $W_B \leq z^2 W_B'$. Since the measurement result of the noise part is determinate, the maximal guessing probability $p_1(a|x)$ for this state is $(1-z) + z\cos^2\theta$. From our derivation process above, the relationship between W_B' and θ is $W_B' \leq \sin 2\theta$. From all these conditions, the $\max p_1(a|x)$ for state (8.37) is

$$\max p_1(a|x) \leq 1 - z + \frac{z}{2}\left(1 + \sqrt{1 - W_B^2/z^2}\right) \tag{8.38}$$

Moreover, another possible scenario is that Alice and Bob share a Werner state (6.7) between them. In this situation, the dimension witness is $W_B \leq z^2$, the maximal guessing probability is $\max p_2(a|x) \leq 1 - z/2$, then

$$\max p_2(a|x) \leq 1 - \frac{\sqrt{W_B}}{2} \tag{8.39}$$

The three upper bounds in Eqn. (8.36), Eqn. (8.38) and Eqn. (8.39) is shown in Fig. (8.3). From this figure, we can see that the guessing probability in Eqn. (8.36) is the worst-case scenario.

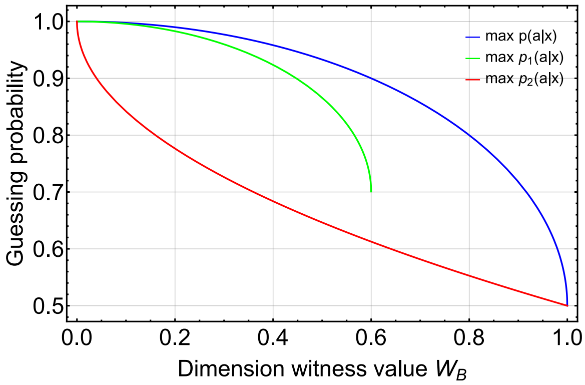


Figure 8.3: The guessing probability upper bounds in three different scenarios. This figure is adapted from supplementary of our publication [3].

Combining Eqn. (8.31) and Eqn. (8.36), the upper bound of $p_g(ab|xy)$ under the SDI conditions can be obtained

$$\begin{aligned} p_g(ab|xy) &\leq \max_{a,x} p(a|x)\frac{1}{2}\sum_y \max_{x,a,b} p(b|(x,a),y) \\ &\leq \left(\frac{1+\sqrt{1-W_B^2}}{2}\right)\frac{1}{2}\left(1+\sqrt{\frac{1+\sqrt{1-W_B^2}}{2}}\right) \end{aligned} \tag{8.40}$$

This guessing probability is derived from W_B. From W_A, a similar upper bound can be derived. Since we define $W_{\text{rsp}} = \min\{W_A, W_B\}$, the larger guessing probability between W_A and W_B will be chosen as $p_g(ab|xy)$, which means the upper bound of $p_g(ab|xy)$ is

$$p_g(ab|xy) \leq \left(\frac{1+\sqrt{1-W_{\text{rsp}}^2}}{2}\right)\frac{1}{2}\left(1+\sqrt{\frac{1+\sqrt{1-W_{\text{rsp}}^2}}{2}}\right) \tag{8.41}$$

As we can see, guessing probability $p_{\text{guess}}(ab|xy)$ from W_{rsp} is not the same as the one from [4]. The difference is caused by $\max_{x,a} p(a|x)$, which represents the quantum measurement from the state preparation process.

Appendix for Chapter 7

Guessing probability of $\mathrm{p(ab|xy)}$ in chapter 7

We show here how to get the guessing probability mentioned in the Eqn. 7.3 of chapter 7. We define the guessing probability of the results of one specific prepare-and-measure run as

$$p_{x,y,\lambda,\mu} = \max_b p(b|x,y,\lambda,\mu) \tag{8.42}$$

where λ, μ represent the internal states of state preparation and measurement devices, and they are independent of each other. Since we are using uniformly distributed input a and b, the guessing probability p_{guess} of all the prepare-and-measure runs is defined as the average of all the combination of different states and measurements:

$$p_{\text{guess}} = \frac{1}{18}\sum_{x,y} \max_b p(b|x,y,\lambda,\mu) \tag{8.43}$$

The coefficient $1/18$ is from the fact that we have six states in preparation, $a = 0, 1, 2, 3, 4, 5$, and three measurement bases $b = 0, 1, 2$. They have 18 different combinations, and each combination will appear with probability $1/18$. Next, we provide an upper bound for this guessing probability. Obviously, we have

$$\begin{aligned} &\frac{1}{18}\sum_{x,y}\max_b p(b|x,y,\lambda,\mu) \\ &\le \frac{1}{3}\max_x \sum_y \max_b p(b|x,y,\lambda,\mu). \end{aligned} \tag{8.44}$$

The "$\le$" in this equation is because we choose one state from the six preparation states, which leads to the maximum guessing probability over the three possible measurement bases. The upper bound of $\frac{1}{3}\max_x \sum_y \max_b p(b|x,y,\lambda,\mu)$, is

$$\begin{aligned} &\frac{1}{3}\max_x \sum_y \max_b p(b|x,y,\lambda,\mu) \\ &\le \frac{1}{6}(3+\sqrt{3}\sqrt{1+2\cos\theta}) \end{aligned} \tag{8.45}$$

where $0 \le \theta \le \pi/2$ is the angle among the three measurement bases as shown in Fig. 8.4

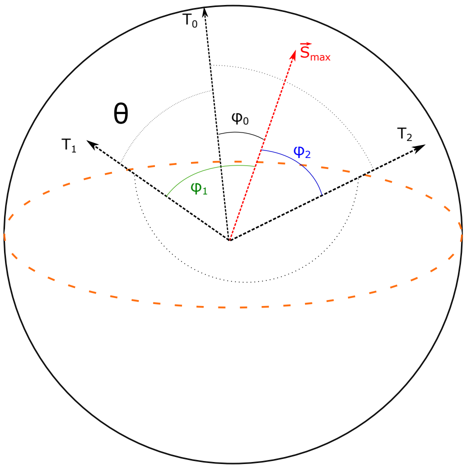


Figure 8.4: One state with three measurement settings in a Bloch sphere. The angel between T_0 and T_1, T_1 and T_2, T_0 and T_2 are all θ (this scenario maximize the guessing probability and will be proved in the following content). The state $\vec{S}_{\max}$ is from the six preparation states which can maximize the guessing probability among T_0, T_1 and T_2. The probabilities of the outcome, say, $b = 1$ is given by projections of state $\vec{S}_{\max}$ onto $T_{0,1,2}$. The maximum average probability of all the three measurements can only be obtained when $\phi_0 = \phi_1 = \phi_2 = 2\cos^{-1}\left(\frac{\sqrt{\sqrt{6}\cos\theta+3+3}}{\sqrt{6}}\right)$, which will be derived shortly.

Next, we need to connect this upper bound with the value of dimension witness W_3. According to Eqn.(14) in [83], W_3 can be represented as:

$$
\begin{aligned}
W_3 = &\begin{vmatrix} \vec{S}_{01}\cdot\vec{T}_0 & \vec{S}_{23}\cdot\vec{T}_0 & \vec{S}_{45}\cdot\vec{T}_0 \\ \vec{S}_{01}\cdot\vec{T}_1 & \vec{S}_{23}\cdot\vec{T}_1 & \vec{S}_{45}\cdot\vec{T}_1 \\ \vec{S}_{01}\cdot\vec{T}_2 & \vec{S}_{23}\cdot\vec{T}_2 & \vec{S}_{45}\cdot\vec{T}_2 \end{vmatrix} \\
=&\vec{S}_{01}\cdot\vec{T}_0\left|\left(\vec{S}_{23}\times\vec{S}_{45}\right)\cdot\left(\vec{T}_1\times\vec{T}_2\right)\right| - \vec{S}_{23}\cdot\vec{T}_0\left|\left(\vec{S}_{01}\times\vec{S}_{45}\right)\cdot\left(\vec{T}_1\times\vec{T}_2\right)\right| + \\
&\vec{S}_{45}\cdot\vec{T}_0\left|\left(\vec{S}_{01}\times\vec{S}_{23}\right)\cdot\left(\vec{T}_1\times\vec{T}_2\right)\right| \\
\leq&\vec{S}_{01}\cdot\vec{T}_0\left|\left(\vec{S}_{23}\times\vec{S}_{45}\right)\cdot\left(\vec{T}_1\times\vec{T}_2\right)\right| + \vec{S}_{23}\cdot\vec{T}_0\left|\left(\vec{S}_{01}\times\vec{S}_{45}\right)\cdot\left(\vec{T}_1\times\vec{T}_2\right)\right| + \\
&\vec{S}_{45}\cdot\vec{T}_0\left|\left(\vec{S}_{01}\times\vec{S}_{23}\right)\cdot\left(\vec{T}_1\times\vec{T}_2\right)\right| \\
\leq&\vec{S}_{01}\cdot\vec{T}_0\left|\vec{S}_{23}\times\vec{S}_{45}\right|\cdot\left|\vec{T}_1\times\vec{T}_2\right| + \vec{S}_{23}\cdot\vec{T}_0\left|\vec{S}_{01}\times\vec{S}_{45}\right|\cdot\left|\vec{T}_1\times\vec{T}_2\right| + \\
&\vec{S}_{45}\cdot\vec{T}_0\left|\vec{S}_{01}\times\vec{S}_{23}\right|\cdot\left|\vec{T}_1\times\vec{T}_2\right| \\
=&\left(\vec{S}_{01}\cdot\left|\vec{S}_{23}\times\vec{S}_{45}\right| + \vec{S}_{23}\cdot\left|\vec{S}_{01}\times\vec{S}_{45}\right| + \vec{S}_{45}\cdot\left|\vec{S}_{01}\times\vec{S}_{23}\right|\right)\cdot\left(\vec{T}_0\cdot\left|\vec{T}_1\times\vec{T}_2\right|\right).
\end{aligned}
\tag{8.46}
$$

Using the properties of scalar triple product $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})$ and $\mathbf{a}\times\mathbf{b} = -(\mathbf{b}\times\mathbf{a})$ (where $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are vectors), the above equation can be further written as

$$
W_3 \leq \left|\left(\vec{S}_{01}\times\vec{S}_{23}\right)\cdot\vec{S}_{45}\right|\cdot\left|\left(\vec{T}_0\times\vec{T}_1\right)\cdot\vec{T}_2\right| \leq \left|\left(\vec{T}_0\times\vec{T}_1\right)\cdot\vec{T}_2\right|. \tag{8.47}
$$

Notice that the angle among the three Bloch vectors $T_{0,1,2}$ is θ. We can represent W_3 as a function of θ:

$$
W_3 \leq 2\sin\left(\frac{\theta}{2}\right)\sin(\theta)\sqrt{1-\frac{1}{2\cos(\theta)+2}} \tag{8.48}
$$

Considering 8.45 and 8.48, the upper bound of p_{guess} can be derived as Eqn. 7.3 in Chapter. 7.

During the deriving of p_{guess}, we have assumed that the angles among three different measurement bases are the same, and also the state $\vec{S}_{\text{max}}$ (see Fig. 8.5) has the same angle towards them. In a real experiment, the angles among three different measurement bases can have three different values, and we prove that only when these three different angles have the same value, the upper bound of p_{guess} will be maximized with a given dimension witness value W_3. Suppose the angles among three measurement bases are different from each other, as shown in Fig 8.5.

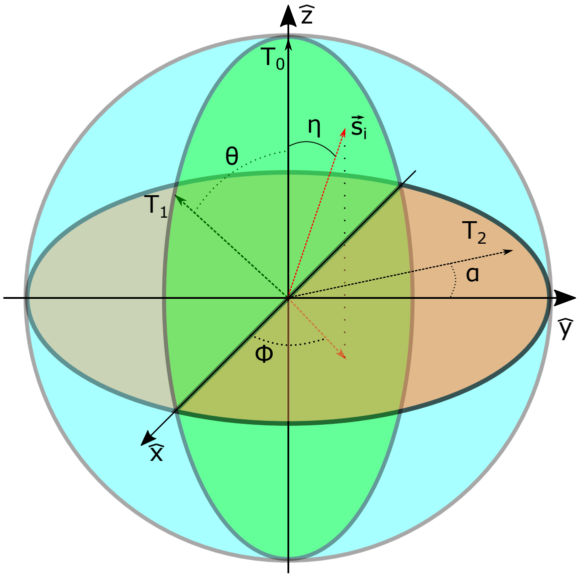


Figure 8.5: Angles between different measurement bases defined in a Bloch sphere. In this figure, measurement basis T_0 is overlapped with $\hat{z}$ axis, T_1 is in the $\hat{x}\hat{z}$ plane, and T_2 has an angle α towards $\hat{y}$ axis. The state $\vec{S}_{\text{max}}$ is from the six preparation states which can maximize the guessing probability among T_0, T_1 and T_2

From Fig 8.5 we can see that the angle between T_0 and T_1 is θ, the angle between T_2 and $\hat{y}$ axis is α. We assume the six preparation states have the following relationship: $\vec{S}_0 = -\vec{S}_1$, $\vec{S}_2 = -\vec{S}_3$, $\vec{S}_4 = -\vec{S}_5$, and $\vec{S}_0, \vec{S}_2, \vec{S}_4$ are mutually perpendicular to each other (one special case is $\vec{S}_0 = -\vec{S}_1 = \hat{z}$, $\vec{S}_2 = -\vec{S}_3 = \hat{x}$ and $\vec{S}_4 = -\vec{S}_5 = \hat{y}$). This condition guarantees that the six preparation states are in an ideal case, and the value of W_3 is only affected by the angles among different measurement bases T_0, T_1, and T_2. Considering the angles among the three measurement bases and Eqn. 8.47, the dimension witness value W_3 can be derived as

$$W_3 \leq \left| \left(\vec{T_0} \times \vec{T_1} \right) \cdot \vec{T_2} \right| = \sin\theta \cos\alpha. \tag{8.49}$$

It can be easily verified that if we rotate T_2 around $\hat{y}$, the value of W_3 will remain unchanged, so W_3 can be determined by two angles θ and α.

The state $\vec{S}_{\mathrm{max}}$ in the Bloch-sphere is defined by two angles η and ϕ. When measuring this state with T_0 the guessing probability of the result is $\cos^2 \frac{\eta}{2}$, for T_1 the guessing probability is $\cos^2 \frac{\delta}{2}$, where δ is the angle between $\vec{S}_{\mathrm{max}}$ and T_1, and it is derived as

$$\delta = \cos^{-1}(\sin(\eta)\sin(\theta)\cos(\phi) + \cos(\eta)\cos(\theta)). \tag{8.50}$$

And when measuring this state with T_2, the guessing probability is $\cos^2 \frac{\gamma}{2}$, where γ is the angle between T_2 and $\vec{S}_{\mathrm{max}}$, γ is derived as

$$\gamma \geq \left(\frac{\cos^{-1}(\sin(\eta)\sin(\phi)) - \alpha}{2} \right). \tag{8.51}$$

Since T_2 can rotate freely around $\hat{y}$ without changing the value of W_3, when rotating T_2 to the position between $\vec{S}_{\mathrm{max}}$ and $\hat{y}$, and in the same plane formed by $\vec{S}_{\mathrm{max}}$ and $\hat{y}$, we can get the minimum value of γ, which maximize the guessing probability $\cos^2 \frac{\gamma}{2}$.

Next, by averaging the guessing probability among the three measurement bases, we have

$$\begin{aligned} p_{\mathrm{guess}} \leq & \frac{\cos^2 \frac{\eta}{2} + \cos^2 \frac{\delta}{2} + \cos^2 \frac{\gamma}{2}}{3} \\ = & f(\theta, \alpha, \eta, \phi), \end{aligned} \tag{8.52}$$

where $f(\theta, \alpha, \eta, \phi)$ means p_{guess} is determined by four parameters $\theta, \alpha, \eta, \phi$. With a given W_3 value, α can be decided by θ, so p_{guess} can be resolved by three different parameters: θ, η and ϕ. By using the method of Lagrange multipliers, we can get the maximum p_{guess} over these three parameters with a given W_3. When the following conditions are satisfied, p_{guess} can be maximized:

$$\begin{aligned} \eta &= 2\cos^{-1}\left(\frac{\sqrt{\sqrt{6\cos\theta + 3} + 3}}{\sqrt{6}} \right) \\ \phi &= \cos^{-1}\left(\tan\left(\frac{\theta}{2}\right) \cot\left(2\cos^{-1}\left(\frac{\sqrt{\sqrt{6\sqrt{\cos^2(\theta)} + 3} + 3}}{\sqrt{6}} \right)\right)\right) \end{aligned} \tag{8.53}$$

where θ is determined by W_3

$$\theta = \sin^{-1}\left(\sqrt{\frac{1}{4}\left(1-2\sqrt[6]{(8W_3^2+1)^3}\cos\left(\frac{1}{3}\left(\tan^{-1}\left(\frac{8W_3\sqrt{(1-W_3^2)^3}}{-8{W_3}^4-20{W_3}^2+1}\right)+\pi\right)\right)\right)}\right) \tag{8.54}$$

With the parameters given in Eqn. 8.53 and 8.54, the angles η, δ, γ in Eqn. 8.52 can be determined by W_3, and as expected, they have the same value as η in Eqn. 8.53. Also, when η, δ, γ have the same value, the angles between three measurement bases T_0, T_1, T_2 are the same, and they all have angle θ between each other.

Then the maximum value of p_{guess} is

$$p_{\text{guess}} = \frac{\left(1+\sqrt{1-\frac{2}{3}\left(1-\sqrt{1-\frac{1}{4}\left(1-2\sqrt[6]{\left(8{W_3}^2+1\right)^3}\cos\left(\frac{1}{3}\left(\tan^{-1}\left(\frac{8W_3\sqrt{\left(1-W_3^2\right)^3}}{-8{W_3}^4-20{W_3}^2+1}\right)+\pi\right)\right)\right)}\right)}\right)}{2}, \tag{8.55}$$

which is exactly the one shown as Eqn. 7.3 in the main text. The proof concludes here.

Determining the thresholds

The single-shot readout results are the number of photons detected in a cycle [190], and these photons have a distribution. Since many photons are detected in a single measurement cycle, it is more efficient to assign a discrete outcome based on a threshold condition [190, 191]. As shown in Fig 8.6, histogram (a1) is the charged state distribution after the laser pumping, and (a2) is the distribution of nuclear spin states after the initialization. (a3) is the photon counts distribution of the final nuclear spin states (which represent the measurement results).

For each histogram, we have to set a threshold to make a decision about the photon counts. The first two thresholds th_1 and th_2 filter out the unnecessary NV^0 state and nuclear spin states $|0_I\rangle, |{+1}_I\rangle$. The third threshold th_3 determines the binary value (i.e., 0 or 1) of the measurement results. To avoid the post-selection of the data, we choose 10% data to get the optimized thresholds.

We have three data sets, and the amount of preliminary events in total is 3188160 bits. We optimize the thresholds of each data set separately. The following analysis and figures are based on the first data set. For the other two data sets, the process is similar.

Each preliminary event contains the photon counts of the single-shot readout results of the NV charge state, initialized nuclear spin state, and the final nuclear spin state. In other words, each final nuclear spin state is from a specific initialized nuclear state in a specific NV charge state. Some final nuclear spin states may come from the undesired initialized nuclear spin states $|0_I\rangle$, $|+1_I\rangle$ or NV^0 state, and these final nuclear spin states must be discarded.

Choose the prepare-and-measure pair $|+\rangle$-$\hat{z}$ as an example. For this prepare-and-measure, the single-shot readout results of its charged state, initialized nuclear spin states, and the final nuclear spin states are shown in Fig 8.6. The three sub-figures in row (a) display the unprocessed data from the experiment. In sub-figure (a1), each peak represents different NV charged states; in (a2), each peak corresponds to different initialized nuclear spin states. (a3) is the final nuclear spin state distribution. The three figures in row (b) explain how the data post-processing works. In (b1), for the charged state decision measurement, the NV center is not only negatively charged, but some of them are also neutrally charged. We only need the negatively charged state NV^-, so we need to set a threshold to get the NV$^-$ state. In (b2), we prepare the nuclear spin into $|-1_\mathrm{I}\rangle$, but not all the nuclear spin can be initialized into this state. Some of them can be initialized into $|0_\mathrm{I}\rangle$ or $|+1_\mathrm{I}\rangle$. So there are two peaks in the distribution of photon counts. Each peak represents different nuclear spin states. As we only need $|-1_\mathrm{I}\rangle$ (which is shown as green parts), we must set thresholds to discard the undesired spin states. After setting the thresholds for the charge states and the initialized nuclear spin states, the photon counts distribution of the final nuclear spin states will also change. With different thresholds in the charged state and the initialized nuclear spin states distribution, the final size of the raw bits also changes accordingly.

When high readout fidelity is preferred, we set the thresholds for the NV$^-$ state and $|-1_\mathrm{I}\rangle$ state from the peaks in the histograms. The threshold of the charged state is 66, and of the initialized nuclear spin state is 46, which is shown as line 1 in sub-figure (b1) and (b2) of Fig 8.6. The raw bits distribution is shown as the lower part of (b3), which is only 3.43% of all the measurement events. The fidelities of preparing NV$^-$ and $|-1_\mathrm{I}\rangle$ with threshold 1 is more than 99%.

When the maximum output randomness is preferred, set both thresholds to line 2, and the distribution of the raw bits is shown in the upper part of (b3). The fidelity of NV$^-$ is about 98%, and $|-1_\text{I}\rangle$ is about 95%. The raw bits size is about 11.83% of all the measurement events.

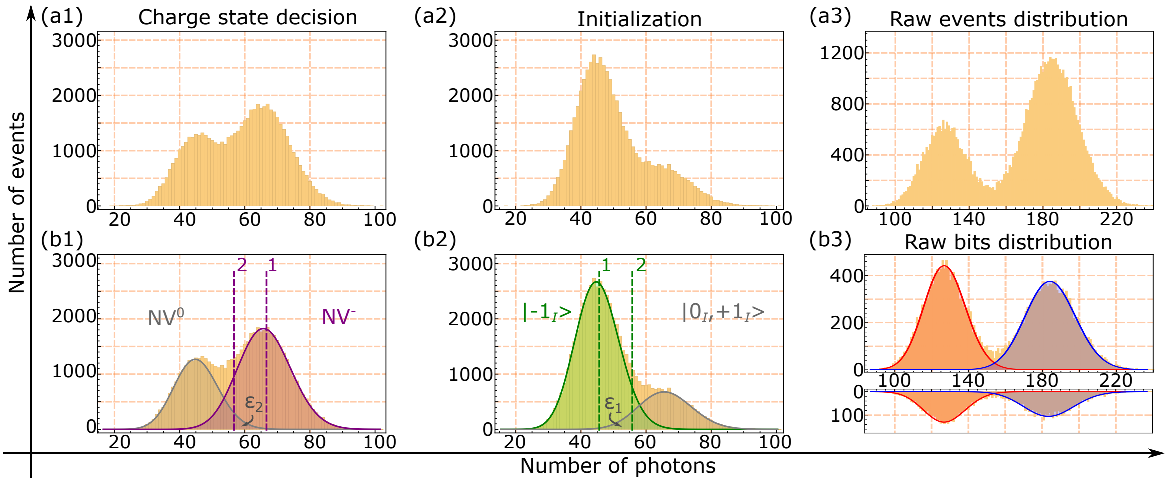


Figure 8.6: The histogram distributions of charge state, nuclear spin initialization and the raw events. In all the sub-figures, the vertical axis is the number of events, and the horizontal axis is the number of photons. The photon counts distribution of the single-shot readout for the charged states, the initialized nuclear spin states, and the final nuclear spin states are shown in row(a). In row(b), we show how the thresholds of nuclear spin initialization and the charged NV states affect the size of the raw bits. The readout error ϵ_1 and ϵ_2 will also change with different thresholds.

We have the raw bits when the thresholds of the charged states and the initialized nuclear spin states are settled. The raw bits have two different peaks, corresponding to two different nuclear spin states. We map the left peak to random bit “0” and the right peak to random bit “1”. To fully distinguish all the raw bits, we need to set a threshold in the distribution of the raw bits. We can either set one threshold or two different thresholds to distinguish between “0” and “1”. Again, we use $|+\rangle$-$\hat{z}$ as an example. Choose the raw bit distribution in the lower part of (b3) in Fig 8.6, and re-show it in Fig 8.7.

In this raw bits distribution, we can set thresholds for each peak from the optimal point, which is shown as the green and blue dashed line in Fig 8.7. In this case, the readout fidelity for “0” and “1” are both higher than 99.999%, but almost half of the raw bits are further discarded. To keep the size of raw bits as large as possible, we can move thresholds from position 1 towards 2 (green dashed curve), where two thresholds are overlapped. In position 2, all the raw bits remain, and the readout fidelities of “0” and “1” are about 99%, which is still very high. So, for the distribution of the raw bits, we set the threshold as the green dashed line.

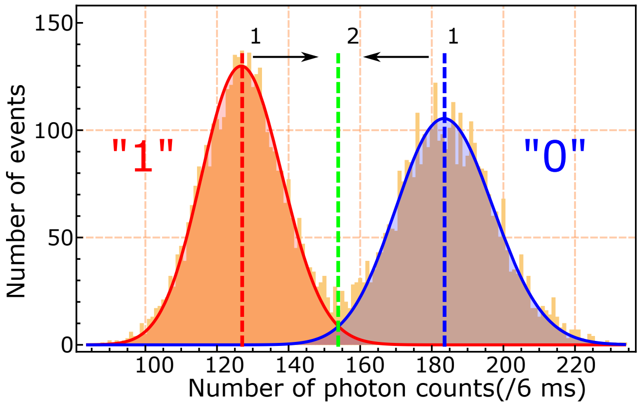


Figure 8.7: The raw bits distribution with different thresholds. When changing the thresholds from threshold 1 (red dashed line in the left and blue dashed line in the right) to threshold 2 (green dashed line in the middle), the average single-shot readout fidelity changes from 99.999% to 99%, the decrease is not so significant, but the size of the raw bits is nearly doubled.

As mentioned before, different thresholds can be settled for the thresholds of the charged states and the initialized nuclear spin states distribution. In our case, we choose the thresholds which can maximize the output randomness. Now we explain it in more detail.

When the thresholds in (b1) and (b2) of Fig 8.6 are moving from position 1 to 2, the raw bits size increases, and also the final extracted randomness reaches the maximum. If we continue to move the thresholds from position 2 to enlarge the size of the raw bits further, the extracted randomness will drop. This is because the total fidelity drops, and so does the dimension witness value. Their relationship is shown in Fig. (8.8).

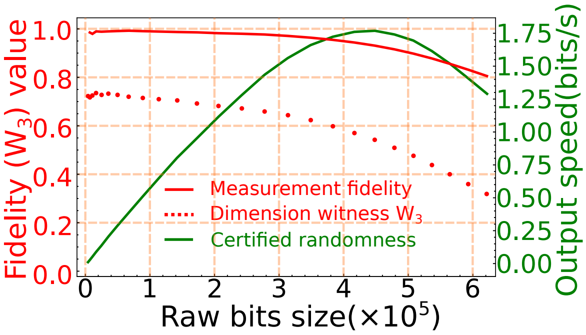


Figure 8.8: The total fidelity versus the randomness output speed. The red curve is the total fidelity, which is affected by the threshold settings. The red dotted curve is the value of W_3 with different preparation fidelities. The green curve is the total randomness output speed with different dimension witness values.

We can see from Fig. (8.8) that the total randomness output is not only influenced by the dimension witness value but also by the size of the raw bits. When we increase the total fidelity of the experiment device, as expected, the value of dimension witness W_3 will also be increased. In the meanwhile, the size of the raw bits is decreasing. The growth of dimension witness value and the drop of raw bits size will make the final output randomness behave as the green curve in this figure.

The drop in the total fidelity and the dimension witness value is mainly due to the threshold settings in figures (b1) and (b2) of Fig. 8.6. When moving thresholds from position 1 towards position 2, the readout error ϵ_1 and ϵ_2 will increase, and the size of the raw bits will also increase. When the thresholds reach position 2, the output randomness reaches the maximum. If we further move the thresholds away from position 2, the size of the raw bits will still increase, but the total fidelity (correspondingly the dimension witness value) will drop faster, so the total extractable randomness will drop. Their relationship is shown in Fig. (8.8).

When all the thresholds are fixed, we now investigate the memory effect of the adjacent experimental runs. The differences between the conditional probabilities $p(1|0)$ and $p(1|1)$ are shown in Fig. 8.9. From this figure, we can see that our experimental setup is memoryless. All the differences can be explained by 3σ shot noise.

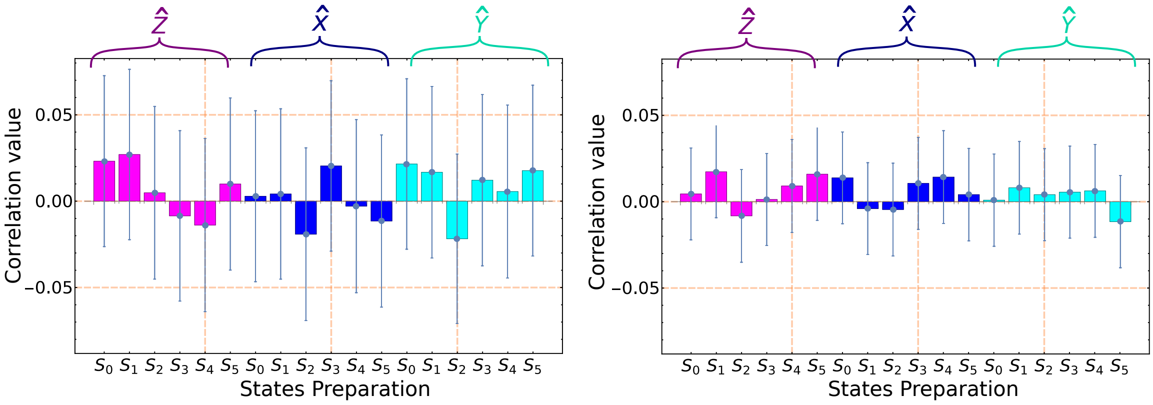


Figure 8.9: The conditional probability differences of 18 different P&M pairs with different threshold settings. The left figure is the differences with 98% total fidelity, and the right figure is the differences with 92% total fidelity (which corresponds to the maximum randomness output). From the figures, we can see that the conditional probability differences do not change so much with the changing of different thresholds. This proves that our experimental device is memoryless with different threshold settings. The differences here are calculated the same way described in the main text.

From the above analysis, we choose to maximize the output randomness. To avoid post-selection of the experimental results, we take 10% measurement events as sample data from each data set to get the thresholds value of the charged state NV^-, the nuclear spin state $|-1_\mathrm{I}\rangle$ initialization, and the raw bits distributions. This sample data for the first data set contains 105,948 measurement events. The photon counts distribution of the sample data, and test data is shown in Fig. 8.10. This figure indicates that the sample data and the rest data have an identical distribution. This not only shows the robustness of our experimental setup but also means that the thresholds from the sample data can be applied to the rest data to maximize the output randomness. After we obtain the value of three thresholds, we get the raw bits with 0 and 1. Then we calculate the dimension witness value and apply our models to certify quantum randomness from the raw bits.

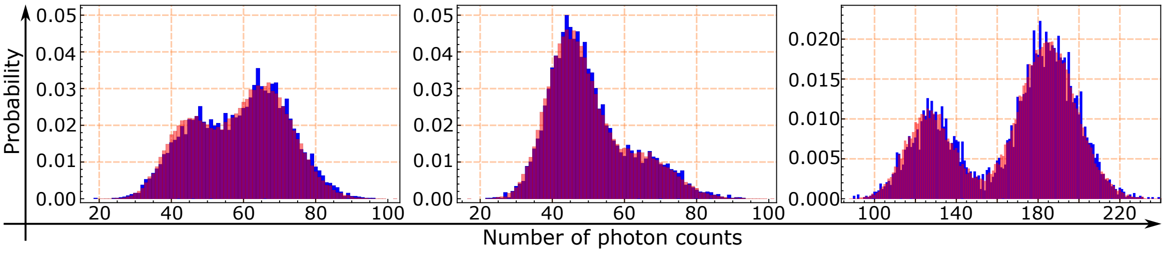


Figure 8.10: The comparison between the photon counts distribution of sample data and the rest valid data. The blue part is the sample data, and the red part is the rest. The overlapping part is purple. From left to right, we can see that the photon counts distribution of charged states, the initialized nuclear spin states, and the final nuclear spin states are identical between the sample data and the test data.

W_2 and W_3 protocols in chapter 4

In the previous analysis, all the threshold settings are optimized for the W_3 model. We prove here that the thresholds also fit the W_2 model. As shown in Fig. 8.11, when the total fidelity is dropping, the value of W_2 and W_3 have similar behaviors (left figure), and also the output speed of the randomness (right figure). So the threshold settings we optimized for the W_3 model are also suitable for the W_2 model.

324 prepare-and-measure pairs

Because of the complexity of the operation, we can only input one sequence with 36 different P&M pairs each time. In this one sequence, if we input pairs randomly, it is infeasible to

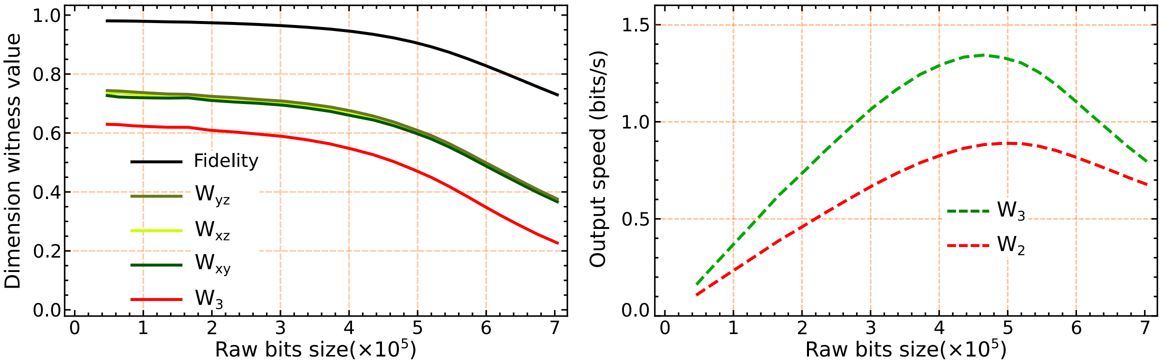


Figure 8.11: The similar behavior of W_2 and W_3 model with different raw bits size. For the figure on the left, from above to bottom, the lines are total fidelity line (which shares the same vertical axis with the dimension witness value), 2-D dimension witness value in yz, xz, xy planes, and the W_3 value. In the right figure, the upper red dashed line is the randomness output speed of the W_2 model, and the lower green dashed line is the output speed of the W_3 model.

implement all 18 possible pairs. This makes it very difficult to investigate the memory effect of each pair. To show that our setup is memoryless for any P&M pair, we specially design a sequence with 324 pairs. During our experiment, we divide these 324 pairs into nine different sections, and each section contains 36 pairs. All 324 pairs can be implemented in every nine input sequences. The 324 pairs are designed in such a way that each of the 18 P&M pairs appears with the same frequency, and after the current pair, all the 18 different pairs appear with the same frequency in the next coming pair. For example, for the P&M pair S_0-$\hat{z}$, the next P&M pair can be any pair from all 18 pairs. In this way, we can analyze the memory effect of S_0-$\hat{z}$ on all the pairs.

The specially designed P&M pairs are shown here.
$\{1x\}\ \{1x\}\ \{2y\}\ \{1x\}\ \{3z\}\ \{1y\}\ \{4y\}\ \{1z\}\ \{5z\}\ \{1x\}\ \{6x\}\ \{2y\}\ \{2x\}\ \{3z\}\ \{2y\}\ \{4y\}$
$\{2z\}\ \{5z\}\ \{2x\}\ \{6x\}\ \{3y\}\ \{3x\}\ \{4z\}\ \{3y\}\ \{5y\}\ \{3z\}\ \{6z\}\ \{4x\}\ \{4x\}\ \{5y\}\ \{4x\}\ \{6z\}$
$\{5y\}\ \{5y\}\ \{6z\}\ \{6z\}\ \{1z\}\ \{1x\}\ \{2x\}\ \{1y\}\ \{3x\}\ \{1z\}\ \{4y\}\ \{1y\}\ \{5z\}\ \{1z\}\ \{6x\}\ \{2x\}$
$\{2y\}\ \{3x\}\ \{2z\}\ \{4y\}\ \{2y\}\ \{5z\}\ \{2z\}\ \{6x\}\ \{3x\}\ \{3y\}\ \{4x\}\ \{3z\}\ \{5y\}\ \{3y\}\ \{6z\}\ \{4z\}$
$\{4x\}\ \{5x\}\ \{4y\}\ \{6x\}\ \{5z\}\ \{5y\}\ \{6y\}\ \{6z\}\ \{1z\}\ \{1z\}\ \{2x\}\ \{1x\}\ \{3y\}\ \{1x\}\ \{4z\}\ \{1y\}$
$\{5y\}\ \{1z\}\ \{6z\}\ \{2x\}\ \{2x\}\ \{3y\}\ \{2x\}\ \{4z\}\ \{2y\}\ \{5y\}\ \{2z\}\ \{6z\}\ \{3x\}\ \{3x\}\ \{4y\}\ \{3x\}$
$\{5z\}\ \{3y\}\ \{6y\}\ \{4z\}\ \{4z\}\ \{5x\}\ \{4x\}\ \{6y\}\ \{5x\}\ \{5z\}\ \{6y\}\ \{6y\}\ \{1y\}\ \{1z\}\ \{2z\}\ \{1x\}$
$\{3x\}\ \{1y\}\ \{4x\}\ \{1z\}\ \{5y\}\ \{1y\}\ \{6z\}\ \{2z\}\ \{2x\}\ \{3x\}\ \{2y\}\ \{4x\}\ \{2z\}\ \{5y\}\ \{2y\}\ \{6z\}$
$\{3z\}\ \{3x\}\ \{4x\}\ \{3y\}\ \{5x\}\ \{3z\}\ \{6y\}\ \{4y\}\ \{4z\}\ \{5z\}\ \{4x\}\ \{6x\}\ \{5y\}\ \{5x\}\ \{6z\}\ \{6y\}$
$\{1y\}\ \{1y\}\ \{2z\}\ \{1z\}\ \{3x\}\ \{1x\}\ \{4y\}\ \{1x\}\ \{5z\}\ \{1y\}\ \{6y\}\ \{2z\}\ \{2z\}\ \{3x\}\ \{2x\}\ \{4y\}$
$\{2x\}\ \{5z\}\ \{2y\}\ \{6y\}\ \{3z\}\ \{3z\}\ \{4x\}\ \{3x\}\ \{5y\}\ \{3x\}\ \{6z\}\ \{4y\}\ \{4y\}\ \{5z\}\ \{4z\}\ \{6x\}$
$\{5x\}\ \{5y\}\ \{6x\}\ \{6z\}\ \{1z\}\ \{1y\}\ \{2y\}\ \{1z\}\ \{3z\}\ \{1x\}\ \{4x\}\ \{1y\}\ \{5x\}\ \{1z\}\ \{6y\}\ \{2y\}$
$\{2z\}\ \{3z\}\ \{2x\}\ \{4x\}\ \{2y\}\ \{5x\}\ \{2z\}\ \{6y\}\ \{3y\}\ \{3z\}\ \{4z\}\ \{3x\}\ \{5x\}\ \{3y\}\ \{6x\}\ \{4z\}$

$\{4y\}$ $\{5y\}$ $\{4z\}$ $\{6z\}$ $\{5x\}$ $\{5x\}$ $\{6y\}$ $\{6x\}$ $\{1x\}$ $\{1z\}$ $\{2y\}$ $\{1y\}$ $\{3z\}$ $\{1z\}$ $\{4x\}$ $\{1x\}$ $\{5y\}$ $\{1x\}$ $\{6z\}$ $\{2y\}$ $\{2y\}$ $\{3z\}$ $\{2z\}$ $\{4x\}$ $\{2x\}$ $\{5y\}$ $\{2x\}$ $\{6z\}$ $\{3y\}$ $\{3y\}$ $\{4z\}$ $\{3z\}$ $\{5x\}$ $\{3x\}$ $\{6y\}$ $\{4x\}$ $\{4z\}$ $\{5y\}$ $\{4y\}$ $\{6z\}$ $\{5z\}$ $\{5x\}$ $\{6x\}$ $\{6y\}$ $\{1y\}$ $\{1x\}$ $\{2z\}$ $\{1y\}$ $\{3y\}$ $\{1z\}$ $\{4z\}$ $\{1x\}$ $\{5x\}$ $\{1y\}$ $\{6x\}$ $\{2z\}$ $\{2y\}$ $\{3y\}$ $\{2z\}$ $\{4z\}$ $\{2x\}$ $\{5x\}$ $\{2y\}$ $\{6x\}$ $\{3z\}$ $\{3y\}$ $\{4y\}$ $\{3z\}$ $\{5z\}$ $\{3x\}$ $\{6x\}$ $\{4y\}$ $\{4x\}$ $\{5z\}$ $\{4y\}$ $\{6y\}$ $\{5z\}$ $\{5z\}$ $\{6x\}$ $\{6x\}$ $\{1x\}$ $\{1y\}$ $\{2x\}$ $\{1z\}$ $\{3y\}$ $\{1y\}$ $\{4z\}$ $\{1z\}$ $\{5x\}$ $\{1x\}$ $\{6y\}$ $\{2x\}$ $\{2z\}$ $\{3y\}$ $\{2y\}$ $\{4z\}$ $\{2z\}$ $\{5x\}$ $\{2x\}$ $\{6y\}$ $\{3x\}$ $\{3z\}$ $\{4y\}$ $\{3y\}$ $\{5z\}$ $\{3z\}$ $\{6x\}$ $\{4x\}$ $\{4y\}$ $\{5x\}$ $\{4z\}$ $\{6y\}$ $\{5y\}$ $\{5z\}$ $\{6z\}$ $\{6x\}$, where $i = \{1, 2, 3, 4, 5, 6\}$, which represents the preparation, and $\{x, y, z\}$ is the measurement bases. Each pair of the 18 pairs appear with frequency 18, and after a specific pair (such as $\{1x\}$), 18 different pairs appear with the same probability.

Selected Mathematical code

Here we provide some of our Mathematica code to show how the experimental data in Chapter 4 and Chapter 6 is post-processed.

Codes for NV based single-photon QRNG

The raw experimental data is stored as $.dat$ files, and in one $.dat$ file, each detection event (one click in the detector) is recorded in a 128 Bit binary format (64-bit, which detector has clicked, and 64-bit with the time tags in ps). The Mathematical code to read the raw data file and plot the anti-bunching curve is shown below.

```
rawdata=BinaryReadList[filepath<>"a62a4f54.dat",{"UnsignedInteger16",
"UnsignedInteger16","UnsignedInteger16","UnsignedInteger16",
"UnsignedInteger64"}];
(*the128bitdataisinformat{0,0,1,0,424254276842064297},where
thethirdnumberiseither0or1 ,meanstheclickintwodifferent
detectors ,andthefifthnumberisthetimetaginps*)
timedelay={};
bitlong=Length[rawdata];
startcounts=rawdata[[1;;bitlong -1,{3,5}]];
endcounts=rawdata[[2;;bitlong ,{3,5}]];
positivedelay=DeleteCases[MapThread[If[#1[[1]]==0&&#2[[1]]==
1,#2[[2]]-#1[[2]]]&,{startcounts ,endcounts}],Null];
negativedelay=DeleteCases[MapThread[If[#1[[1]]==1&&#2[[1]]==
0,#1[[2]]-#2[[2]]]&,{startcounts ,endcounts}],Null];
timedelay=Join[timedelay ,negativedelay ,positivedelay];
tbin=500;(*thetimebinsizeinps*)
delaytime=200000;(*thetimewindowoftheantibunchingcurve ,inps*)
datadelay=BinCounts[timedelay,{-delaytime ,delaytime ,tbin}];
delaycount=MapThread[{#1,#2}&,{Range[-delaytime+tbin/2,
delaytime-tbin/2,tbin],datadelay}];
ListPlot[delaycount ,Joined->True,AxesLabel->
{"DelayedTimeps","Coincidencecounts"},PlotRange->All]
```

With the antibunching curve being obtained, the randomness can be further post-processed with the second model and third model mentioned in Chapter 4.

Codes for Bell test QRNG

For this work, we have in total 55 568 loop-hole free Bell test events. Due to the event-ready scenario, the events from $|\Psi^+\rangle$ state and $|\Psi^-\rangle$ state are mixed in the raw data file. In the next Mathematica codes, we show how the S value for the CHSH inequality is calculated from the total 55 568 events.

```
rawdata=Drop[Import[filepath <>"Events-All.csv";,"Table",
"FieldSeparators"->";"],1];
bellstate= data[[All,2]];(*correspondingBellstate*)
chosenbellstate=DeleteCases[MapThread[If[#1=="Psi+",#2]&,
{bellstate,rawdata}],Null];
settingdata=chosenbellstate[[All,{3,5}]];(*measurementsettings*)
resultsdata=chosenbellstate[[All,{4,6}]];(*measurementresults*)
pab[settings_]:=Module[{tempdata,counts},
tempdata=DeleteCases[MapThread[If[#1==settings,#2]&,
{settingdata,resultsdata}],Null];
counts=Length[tempdata];
pabxy=(SequenceCount[tempdata,{{0,0}}]+
SequenceCount[tempdata,{{1,1}}])/counts]
svalue=Abs[2(pab[{0,0}]+pab[{0,1}]+pab[{1,0}]-pab[{1,1}])-2]//N
```

In the above codes, the S value for Bell state $|\Psi^{+}\rangle$ is calculated. For the same Bell state, the Mathematica code to calculate the RSP-dimension witness from the raw data is shown below

```
(*Alice's dimension witness*)
Clear[rawdata,settingdata,resultsdata]
settingdata=rawdata[[All,{3,5}]];(*measurementsetting*)
set00=SequenceCount[settingdata,{{0,0}}];
set01=SequenceCount[settingdata,{{0,1}}];
set10=SequenceCount[settingdata,{{1,0}}];
set11=SequenceCount[settingdata,{{1,1}}];
resultsdata=rawdata[[All,{4,6}]];(*measurementresult*)
temppsi1=DeleteCases[MapThread[If[#1==0(*Bsetting*)&&#2==1(*Bresult*)
&&#3==0(*Asetting*),#4(*Aresult*)]&,{settingdata[[All,2]],
resultsdata[[All,2]],settingdata[[All,1]],resultsdata[[All,1]]}],Null];
temppsi2=DeleteCases[MapThread[If[#1==0(*Bsetting*)&&#2==0(*Bresult*)
&&#3==0(*Asetting*),#4(*Aresult*)]&,{settingdata[[All,2]],
resultsdata[[All,2]],settingdata[[All,1]],resultsdata[[All,1]]}],Null];
w00=Total[temppsi1]/Length[temppsi1]-Total[temppsi2]/Length[temppsi2];
(*w00=p(0,0)-p(1,0)*)
(*similarly, w01(p(2,0)-p(3,0)),w10(p(0,1)-p(1,1)),w11(p(2,1)-p(3,1))
can be calculated*)
walice=Abs[Det[({{w00,w01},{w10,w11}})]]
```

List of Figures

List of Tables

List of Publications

- X. Chen, J. N. Greiner, J. Wrachtrup, and I. Gerhardt. single-photon randomness based on a defect center in diamond. *Scientific Reports*, 9(1):18474, 2019.

- X. Chen, K. Redeker, R. Garthoff, W. Rosenfeld, J. Wrachtrup, and I. Gerhardt. Certified randomness from a remote-state-preparation dimension witness. *Phys. Rev. A*, 103:042211, Apr 2021.

- X. Chen, M. Kwon, V. Vorobyov, J. Wrachtrup, and I. Gerhardt. Self-testing randomness from a nuclear spin system.(Manuscript in preparation).

References

[1] V. Jacques, E. Wu, F. Grosshans, F. Treussart, P. Grangier, A. Aspect, and J.-F. Roch, "Experimental realization of wheeler's delayed-choice gedanken experiment," *Science*, vol. 315, no. 5814, pp. 966–968, 2007.

[2] J. S. Bell, "On the einstein podolsky rosen paradox," *Physics Physique Fizika*, vol. 1, pp. 195–200, Nov 1964.

[3] X. Chen, K. Redeker, R. Garthoff, W. Rosenfeld, J. Wrachtrup, and I. Gerhardt, "Certified randomness from a remote-state-preparation dimension witness," *Phys. Rev. A*, vol. 103, p. 042211, Apr 2021.

[4] T. Lunghi, J. B. Brask, C. C. W. Lim, Q. Lavigne, J. Bowles, A. Martin, H. Zbinden, and N. Brunner, "Self-testing quantum random number generator," *Phys. Rev. Lett.*, vol. 114, p. 150501, Apr 2015.

[5] J. Moreh, "Randomness, game theory and free will," *Erkenntnis*, vol. 41, no. 1, pp. 49–64, 1994.

[6] L. E. Szabó, "Is quantum mechanics compatible with a deterministic universe? two interpretations of quantum probabilities," *Foundations of Physics Letters*, vol. 8, no. 5, pp. 417–436, 1995.

[7] G. Brassard and P. Raymond-Robichaud, "Can free will emerge from determinism in quantum theory?," in *Is science compatible with free will?*, pp. 41–61, Springer, 2013.

[8] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *Theoretical Computer Science*, vol. 560, pp. 7 – 11, 2014. Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84.

[9] J. G. Rarity, P. C. M. Owens, and P. R. Tapster, "Quantum random-number generation and key sharing," *Journal of Modern Optics*, vol. 41, no. 12, pp. 2435–2444, 1994.

[10] S. J. Lomonaco, “A quick glance at quantum cryptography,” *Cryptologia*, vol. 23, no. 1, pp. 1–41, 1999.

[11] M. Stipčević, “Quantum random number generators and their use in cryptography,” *ArXiv e-prints*, Mar. 2011.

[12] R. P. Carlisle, *Encyclopedia of play in today’s society*, vol. 1. Sage, 2009.

[13] B. Bodhi, *The Discourse on the All-Embracing Net of Views: The Brahmajāla Sutta and its Commentaries*, vol. 209. Buddhist Publication Society, 2007.

[14] D. Matz, *Daily life of the Ancient Romans*. Greenwood Publishing Group, 2002.

[15] “MS Windows NT dice in ancient roman world.” `https://imperiumromanum.pl/en/article/dice-in-ancient-roman-world/`.

[16] J. Sachs *et al.*, *Aristotle’s physics: A guided study*. Rutgers University Press, 1995.

[17] J. E. Lightner, “A brief look at the history of probability and statistics,” *The Mathematics Teacher*, vol. 84, no. 8, pp. 623–630, 1991.

[18] I. Schneider, “Abraham de moivre, the doctrine of chances (1718, 1738, 1756),” in *Landmark Writings in Western Mathematics 1640-1940*, pp. 105–120, Elsevier, 2005.

[19] I. Grattan-Guinness, *Landmark writings in Western mathematics 1640-1940*. Elsevier, 2005.

[20] R. Clausius, *Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie: vorgetragen in der naturforsch. Gesellschaft den 24. April 1865*. éditeur inconnu, 1865.

[21] C. E. Shannon, “A mathematical theory of communication,” *ACM SIGMOBILE mobile computing and communications review*, vol. 5, no. 1, pp. 3–55, 2001.

[22] T. E. Hull and A. R. Dobell, “Random number generators,” *SIAM review*, vol. 4, no. 3, pp. 230–254, 1962.

[23] M. Herrero-Collantes and J. C. Garcia-Escartin, “Quantum random number generators,” *Rev. Mod. Phys.*, vol. 89, p. 015004, Feb 2017.

[24] M. Isida and H. Ikeda, “Random number generator,” *Annals of the Institute of Statistical Mathematics*, vol. 8, no. 2, pp. 119–126, 1956.

[25] J. Manelis, "Generating random noise with radioactive sources," *Electronics (US)*, vol. 34, no. 36, 1961.

[26] C. Vincent, "The generation of truly random binary numbers," *Journal of Physics E: Scientific Instruments*, vol. 3, no. 8, p. 594, 1970.

[27] H. Inoue, H. Kumahora, Y. Yoshizawa, M. Ichimura, and O. Miyatake, "Random numbers generated by a physical device," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 32, no. 2, pp. 115–120, 1983.

[28] G. M. Morris, "Optical computing by monte carlo methods," *Optical Engineering*, vol. 24, no. 1, pp. 86–90, 1985.

[29] D. P. Kroese, T. Brereton, T. Taimre, and Z. I. Botev, "Why the monte carlo method is so important today," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 6, no. 6, pp. 386–392, 2014.

[30] D. Rogers, "Fifty years of monte carlo simulations for medical physics," *Physics in Medicine & Biology*, vol. 51, no. 13, p. R287, 2006.

[31] H. MacGillivray, R. Dodd, B. McNally, J. Lightfoot, H. Corwin, and S. Heathcote, "Monte-carlo simulations of galaxy systems," *Astrophysics and Space Science*, vol. 81, no. 1, pp. 231–250, 1982.

[32] J. C. Spall, *Introduction to stochastic search and optimization: estimation, simulation, and control.* John Wiley & Sons, 2005.

[33] L. Elwart, N. Emerson, C. Enders, D. Fumia, and K. Murphy, "Increasing access to restraining orders for low-income victims of domestic violence: A cost-benefit analysis of the proposed domestic abuse grant program," *Madison, State Bar Association of Wisconsin*, 2006.

[34] K. A. Zimmermann, "History of computers: A brief timeline," *Live science*, 2017.

[35] M. Castells, "The information age: Economy, society and culture (3 volumes)," *Blackwell, Oxford*, vol. 1997, p. 1998, 1996.

[36] P. Voigt and A. Von dem Bussche, "The eu general data protection regulation (gdpr)," *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, vol. 10, no. 3152676, pp. 10–5555, 2017.

[37] E. C. Reinke, “Classical cryptography,” *The Classical Journal*, vol. 58, no. 3, pp. 113–121, 1962.

[38] R. L. Rivest, A. Shamir, and L. Adleman, “A method for obtaining digital signatures and public-key cryptosystems,” *Communications of the ACM*, vol. 21, no. 2, pp. 120–126, 1978.

[39] E. Barker and Q. Dang, “Nist special publication 800-57 part 1, revision 4,” *NIST, Tech. Rep*, vol. 16, 2016.

[40] H. B. Westlund, “Nist reports measurable success of advanced encryption standard,” *Journal of Research of the National Institute of Standards and Technology*, vol. 107, no. 3, p. 307, 2002.

[41] A. K. Ekert, “Quantum cryptography based on bell’s theorem,” *Phys. Rev. Lett.*, vol. 67, pp. 661–663, Aug 1991.

[42] C. H. Bennett, G. Brassard, and N. D. Mermin, “Quantum cryptography without bell’s theorem,” *Phys. Rev. Lett.*, vol. 68, pp. 557–559, Feb 1992.

[43] Y. Mu, J. Seberry, and Y. Zheng, “Shared cryptographic bits via quantized quadrature phase amplitudes of light,” *Optics Communications*, vol. 123, no. 1, pp. 344–352, 1996.

[44] J. Von Neumann, “13. various techniques used in connection with random digits,” *Appl. Math Ser*, vol. 12, no. 36-38, p. 3, 1951.

[45] J. Von Neumann, “Various techniques used in connection with random digits,” *John von Neumann, Collected Works*, vol. 5, pp. 768–770, 1963.

[46] F. James, “A review of pseudorandom number generators,” *Computer physics communications*, vol. 60, no. 3, pp. 329–344, 1990.

[47] M. Matsumoto and T. Nishimura, “Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator,” *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 8, no. 1, pp. 3–30, 1998.

[48] I. Goldberg and D. Wagner, “Randomness and the netscape browser,” *Dr Dobb’s Journal-Software Tools for the Professional Programmer*, vol. 21, no. 1, pp. 66–71, 1996.

[49] D. Ahmad, “Two years of broken crypto: debian’s dress rehearsal for a global pki compromise,” *IEEE Security & Privacy*, vol. 6, no. 5, pp. 70–73, 2008.

[50] P. Ducklin, "Android random number flaw implicated in bitcoin thefts," *Naked Security*, 2013.

[51] A. M. Ferrenberg, D. P. Landau, and Y. J. Wong, "Monte carlo simulations: Hidden errors from "good" random number generators," *Phys. Rev. Lett.*, vol. 69, pp. 3382–3384, Dec 1992.

[52] S. H. Kellert, *In the wake of chaos: Unpredictable order in dynamical systems*. University of Chicago press, 1993.

[53] L. D. Landau and E. M. Lifshitz, *Quantum mechanics: non-relativistic theory*, vol. 3. Elsevier, 2013.

[54] W. Gerlach and O. Stern, "Der experimentelle nachweis der richtungsquantelung im magnetfeld," *Zeitschrift für Physik*, vol. 9, no. 1, pp. 349–352, 1922.

[55] M. Tegmark, "The interpretation of quantum mechanics: Many worlds or many words?," *Fortschritte der Physik: Progress of Physics*, vol. 46, no. 6-8, pp. 855–862, 1998.

[56] N. Wolchover *et al.*, "The universe is as spooky as einstein thought," *Atlantic*, 2017.

[57] L. E. Ballentine, "The statistical interpretation of quantum mechanics," *Reviews of modern physics*, vol. 42, no. 4, p. 358, 1970.

[58] W. Thomas, *STACS 2007: 24th Annual Symposium on Theoretical Aspects of Computer Science, Aachen, Germany, February 22-24, 2007, Proceedings*, vol. 4393. Springer Science & Business Media, 2007.

[59] AndrewRukhin, JuanSoto, JamesNechvatal, M. Smid, ElaineBarker, S. Leigh, MarkLevenson, M. Vangel, DavidBanks, AlanHeckert, JamesDray, and SanVo, *NIST Special Publication 800-22: A Statistical Test Suite for the Validation of Random Number Generators and Pseudo Random Number Generators for Cryptographic Applications*. 04 2010.

[60] S. Pironio, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani, "Device-independent quantum key distribution secure against collective attacks," *New Journal of Physics*, vol. 11, p. 045021, apr 2009.

[61] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and

R. Hanson, "Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres," *Nature*, vol. 526, no. 7575, pp. 682–686, 2015.

[62] W. Rosenfeld, D. Burchardt, R. Garthoff, K. Redeker, N. Ortegel, M. Rau, and H. Weinfurter, "Event-ready bell test using entangled atoms simultaneously closing detection and locality loopholes," *Phys. Rev. Lett.*, vol. 119, p. 010402, Jul 2017.

[63] M. Li, P. Vitányi, *et al.*, *An introduction to Kolmogorov complexity and its applications*, vol. 3. Springer, 2008.

[64] A. A. Abbott, "Quantum random numbers: Certification and generation," Master's thesis, The University of Auckland, 2011.

[65] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, "Quantum random number generation," *npj Quantum Information*, vol. 2, p. 16021, Jun 2016.

[66] A. Acín and L. Masanes, "Certified randomness in quantum physics," *Nature*, vol. 540, no. 7632, pp. 213–219, 2016.

[67] B. Qi, Y.-M. Chi, H.-K. Lo, and L. Qian, "High-speed quantum random number generation by measuring phase noise of a single-mode laser," *Opt. Lett.*, vol. 35, pp. 312–314, Feb 2010.

[68] C. Gabriel, C. Wittmann, D. Sych, R. Dong, W. Mauerer, U. L. Andersen, C. Marquardt, and G. Leuchs, "A generator for unique quantum random numbers based on vacuum states," *Nature Photonics*, vol. 4, no. 10, pp. 711–715, 2010.

[69] T. Symul, S. M. Assad, and P. K. Lam, "Real time demonstration of high bitrate quantum random number generation with coherent laser light," *Appl. Phys. Lett.*, vol. 98, 2011.

[70] Z. Zheng, Y. Zhang, W. Huang, S. Yu, and H. Guo, "6 gbps real-time optical quantum random number generator based on vacuum fluctuation," *Review of Scientific Instruments*, vol. 90, no. 4, p. 043105, 2019.

[71] T. Gehring, C. Lupo, A. Kordts, D. Solar Nikolic, N. Jain, T. Rydberg, T. B. Pedersen, S. Pirandola, and U. L. Andersen, "Homodyne-based quantum random number generator at 2.9 gbps secure against quantum side-information," *Nature communications*, vol. 12, no. 1, pp. 1–11, 2021.

[72] M. Petrov, I. Radchenko, D. Steiger, R. Renner, M. Troyer, and V. Makarov, "Independent security analysis of a commercial quantum random number generator," *arXiv preprint arXiv:2004.04996*, 2020.

[73] J. Clauser, M. Horne, A. Shimony, and R. Holt, "Proposed experiment to test local hidden-variable theories," *Phys. Rev. Lett.*, vol. 23, pp. 880–884, 10 1969.

[74] R. Arnon-Friedman, F. Dupuis, O. Fawzi, R. Renner, and T. Vidick, "Practical device-independent quantum cryptography via entropy accumulation," *Nature Communications*, vol. 9, p. 459, Jan 2018.

[75] P. Bierhorst, E. Knill, S. Glancy, Y. Zhang, A. Mink, S. Jordan, A. Rommal, Y.-K. Liu, B. Christensen, S. W. Nam, M. J. Stevens, and L. K. Shalm, "Experimentally generated randomness certified by the impossibility of superluminal signals," *Nature*, vol. 556, no. 7700, pp. 223–226, 2018.

[76] Y. Liu, Q. Zhao, M.-H. Li, J.-Y. Guan, Y. Zhang, B. Bai, W. Zhang, W.-Z. Liu, C. Wu, X. Yuan, H. Li, W. J. Munro, Z. Wang, L. You, J. Zhang, X. Ma, J. Fan, Q. Zhang, and J.-W. Pan, "Device-independent quantum random-number generation," *Nature*, vol. 562, no. 7728, pp. 548–551, 2018.

[77] E. Knill, Y. Zhang, and P. Bierhorst, "Generation of quantum randomness by probability estimation with classical side information," *Physical Review Research*, vol. 2, Sep 2020.

[78] Y. Zhang, H. Fu, and E. Knill, "Efficient randomness certification by quantum probability estimation," *Phys. Rev. Research*, vol. 2, p. 013016, Jan 2020.

[79] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-A. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger, "Significant-loophole-free test of bell's theorem with entangled photons," *Phys. Rev. Lett.*, vol. 115, p. 250401, Dec 2015.

[80] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer, C. Hodge, A. E. Lita, V. B. Verma, C. Lambrocco, E. Tortorici, A. L. Migdall, Y. Zhang, D. R. Kumor, W. H. Farr, F. Marsili, M. D. Shaw, J. A. Stern, C. Abellán, W. Amaya, V. Pruneri, T. Jennewein, M. W. Mitchell, P. G. Kwiat, J. C. Bienfang, R. P. Mirin, E. Knill, and S. W. Nam, "Strong loophole-free test of local realism," *Phys. Rev. Lett.*, vol. 115, p. 250402, Dec 2015.

[81] J.-Å. Larsson, "Loopholes in bell inequality tests of local realism," *Journal of Physics A: Mathematical and Theoretical*, vol. 47, p. 424003, oct 2014.

[82] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, "Random numbers certified by bell's theorem," *Nature*, vol. 464, pp. 1021 EP –, Apr 2010.

[83] J. Bowles, M. T. Quintino, and N. Brunner, "Certifying the dimension of classical and quantum systems in a prepare-and-measure scenario with independent devices," *Phys. Rev. Lett.*, vol. 112, p. 140407, Apr 2014.

[84] M. Jerger, Y. Reshitnyk, M. Oppliger, A. Potočnik, M. Mondal, A. Wallraff, K. Goodenough, S. Wehner, K. Juliusson, N. K. Langford, and A. Fedorov, "Contextuality without nonlocality in a superconducting quantum system," *Nature Communications*, vol. 7, Oct. 2016.

[85] S. Kochen and E. P. Specker, "The problem of hidden variables in quantum mechanics," *Journal of Mathematics and Mechanics*, vol. 17, no. 1, pp. 59–87, 1967.

[86] M. Um, X. Zhang, J. Zhang, Y. Wang, S. Yangchao, D. L. Deng, L.-M. Duan, and K. Kim, "Experimental Certification of Random Numbers via Quantum Contextuality," *Scientific Reports*, vol. 3, Apr. 2013.

[87] C. Kurtsiefer, S. Mayer, P. Zarda, and H. Weinfurter, "Stable solid-state source of single-photons," *Phys. Rev. Lett.*, vol. 85, pp. 290–293, Jul 2000.

[88] J. F. Clauser and M. A. Horne, "Experimental consequences of objective local theories," *Phys. Rev. D*, vol. 10, pp. 526–535, Jul 1974.

[89] D. Orsucci, J.-D. Bancal, N. Sangouard, and P. Sekatski, "How post-selection affects device-independent claims under the fair sampling assumption," *Quantum*, vol. 4, p. 238, 2020.

[90] L. Childress, M. V. G. Dutt, J. M. Taylor, A. S. Zibrov, F. Jelezko, J. Wrachtrup, P. R. Hemmer, and M. D. Lukin, "Coherent dynamics of coupled electron and nuclear spin qubits in diamond," *Science*, vol. 314, no. 5797, pp. 281–285, 2006.

[91] S. Pezzagna and J. Meijer, "Quantum computer based on color centers in diamond," *Applied Physics Reviews*, vol. 8, no. 1, p. 011308, 2021.

[92] G. Waldherr, Y. Wang, S. Zaiser, M. Jamali, T. Schulte-Herbrüggen, H. Abe, T. Ohshima, J. Isoya, J. F. Du, P. Neumann, and J. Wrachtrup, "Quantum error correction in a solid-state hybrid spin register," *Nature*, vol. 506, pp. 204–207, Feb 2014.

[93] V. Vorobyov, S. Zaiser, N. Abt, J. Meinel, D. Dasari, P. Neumann, and J. Wrachtrup, "Quantum fourier transform for nanoscale quantum sensing," *npj Quantum Information*, vol. 7, p. 124, Aug 2021.

[94] T. Jennewein, U. Achleitner, G. Weihs, H. Weinfurter, and A. Zeilinger, "A fast and compact quantum random number generator," *Review of Scientific Instruments*, vol. 71, no. 4, pp. 1675–1680, 2000.

[95] M. A. Wayne, E. R. Jeffrey, G. M. Akselrod, and P. G. Kwiat, "Photon arrival time quantum random number generation," *Journal of Modern Optics*, vol. 56, no. 4, pp. 516–522, 2009.

[96] M. Wahl, M. Leifgen, M. Berlin, T. Röhlicke, H.-J. Rahn, and O. Benson, "An ultrafast quantum random number generator with provably bounded output bias based on photon arrival time measurements," *Applied Physics Letters*, vol. 98, no. 17, pp. –, 2011.

[97] M. Fürst, H. Weier, S. Nauerth, D. G. Marangon, C. Kurtsiefer, and H. Weinfurter, "High speed optical quantum random number generation," *Opt. Express*, vol. 18, pp. 13029–13037, Jun 2010.

[98] M. Ren, E. Wu, Y. Liang, Y. Jian, G. Wu, and H. Zeng, "Quantum random-number generator based on a photon-number-resolving detector," *Phys. Rev. A*, vol. 83, p. 023820, Feb 2011.

[99] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov, "Invited review article: Single-photon sources and detectors," *Review of scientific instruments*, vol. 82, no. 7, p. 071101, 2011.

[100] H. Inamori, N. Lütkenhaus, and D. Mayers, "Unconditional security of practical quantum key distribution," *The European Physical Journal D*, vol. 41, no. 3, pp. 599–627, 2007.

[101] E. Meyer-Scott, C. Silberhorn, and A. Migdall, "Single-photon sources: Approaching the ideal through multiplexing," *Review of Scientific Instruments*, vol. 91, no. 4, p. 041101, 2020.

[102] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov, "Invited review article: Single-photon sources and detectors," *Review of scientific instruments*, vol. 82, no. 7, p. 071101, 2011.

[103] N. Somaschi, V. Giesz, L. De Santis, J. Loredo, M. P. Almeida, G. Hornecker, S. L. Portalupi, T. Grange, C. Anton, J. Demory, *et al.*, "Near-optimal single-photon sources in the solid state," *Nature Photonics*, vol. 10, no. 5, pp. 340–345, 2016.

[104] S. Thomas and P. Senellart, "The race for the ideal single-photon source is on," *Nature Nanotechnology*, vol. 16, no. 4, pp. 367–368, 2021.

[105] R. J. Glauber, "The quantum theory of optical coherence," *Phys. Rev.*, vol. 130, pp. 2529–2539, Jun 1963.

[106] R. J. Glauber, "Coherent and incoherent states of the radiation field," *Phys. Rev.*, vol. 131, pp. 2766–2788, Sep 1963.

[107] C. Gerry, P. Knight, and P. L. Knight, *Introductory quantum optics*. Cambridge university press, 2005.

[108] R. H. Brown, R. Q. Twiss, and g. surName, "Interferometry of the intensity fluctuations in light-i. basic theory: the correlation between photons in coherent beams of radiation," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 242, no. 1230, pp. 300–324, 1957.

[109] M. Fox, *Quantum Optics: An Introduction.* Oxford University Press, 2006.

[110] A. Gruber, A. Dräbenstedt, C. Tietz, L. Fleury, J. Wrachtrup, and C. v. Borczyskowski, "Scanning confocal optical microscopy and magnetic resonance on single defect centers," *Science*, vol. 276, no. 5321, pp. 2012–2014, 1997.

[111] R. Brouri, A. Beveratos, J.-P. Poizat, and P. Grangier, "Photon antibunching in the fluorescence of individual color centers in diamond," *Opt. Lett.*, vol. 25, pp. 1294–1296, Sep 2000.

[112] M. Jamali, I. Gerhardt, M. Rezai, K. Frenner, H. Fedder, and J. Wrachtrup, "Microscopic diamond solid-immersion-lenses fabricated around single defect centers by focused ion beam milling," *Review of Scientific Instruments*, vol. 85, no. 12, p. 123703, 2014.

[113] A. Beveratos, R. Brouri, T. Gacoin, A. Villing, J.-P. Poizat, and P. Grangier, "single-photon quantum cryptography," *Phys. Rev. Lett.*, vol. 89, p. 187901, Oct 2002.

[114] F. Jelezko, T. Gaebel, I. Popa, M. Domhan, A. Gruber, and J. Wrachtrup, "Observation of coherent oscillation of a single nuclear spin and realization of a two-qubit conditional quantum gate," *Phys. Rev. Lett.*, vol. 93, p. 130501, Sep 2004.

[115] M. W. Doherty, N. B. Manson, P. Delaney, and L. C. L. Hollenberg, "The negatively charged nitrogen-vacancy centre in diamond: the electronic solution," vol. 13, p. 025019, feb 2011.

[116] B. B. Baker and E. T. Copson, *The mathematical theory of Huygens' principle*, vol. 329. American Mathematical Soc, 2003.

[117] T. Young, "Ii. the bakerian lecture. on the theory of light and colours," *Philosophical transactions of the Royal Society of London*, no. 92, pp. 12–48, 1802.

[118] X.-s. Ma, J. Kofler, and A. Zeilinger, "Delayed-choice gedanken experiments and their realizations," *Rev. Mod. Phys.*, vol. 88, p. 015005, Mar 2016.

[119] A. Einstein, "Concerning an heuristic point of view toward the emission and transformation of light," *American Journal of Physics*, vol. 33, no. 5, p. 367, 1965.

[120] L. De Broglie, *Recherches sur la théorie des quanta*. PhD thesis, 1924.

[121] C. Davisson and L. H. Germer, "Diffraction of electrons by a crystal of nickel," *Phys. Rev.*, vol. 30, pp. 705–740, Dec 1927.

[122] I. Estermann and O. Stern, "Beugung von molekularstrahlen," *Zeitschrift für Physik*, vol. 61, no. 1, pp. 95–125, 1930.

[123] T. Rothman, *Everything's Relative: And Other Fables from Science and Technology*. 2003.

[124] C. J. Davisson and L. H. Germer, "Reflection of electrons by a crystal of nickel," *Proceedings of the National Academy of Sciences*, vol. 14, no. 4, pp. 317–322, 1928.

[125] W. Rueckner and J. Peidle, "Young's double-slit experiment with single-photons and quantum eraser," *American Journal of Physics*, vol. 81, no. 12, pp. 951–958, 2013.

[126] S. Eibenberger, S. Gerlich, M. Arndt, M. Mayor, and J. Tüxen, "Matter–wave interference of particles selected from a molecular library with masses exceeding 10 000 amu," *Phys. Chem. Chem. Phys.*, vol. 15, pp. 14696–14700, 2013.

[127] L. Zehnder, "Ein neuer interferenzrefraktor," *Zeitschrift für Instrumentenkunde*, pp. 275–285, 1891.

[128] L. Mach, "Ueber einen interferenzrefrakto," *Zeitschrift für Instrumentenkunde*, pp. 89–93, 1892.

[129] G. Weihs and A. Zeilinger, "Photon statistics at beam-splitters: an essential tool in quantum information and teleportation," *Coherence and Statistics of Photons and Atoms*, pp. 262–288, 2001.

[130] J. A. Wheeler, "The past and the delayed-choice double-slit experiment," in *Mathematical foundations of quantum theory*, pp. 9–48, Elsevier, 1978.

[131] J. A. Wheeler and W. H. Zurek, *Quantum theory and measurement*, vol. 15. Princeton University Press, 2014.

[132] P. D. D. Schwindt, P. G. Kwiat, and B.-G. Englert, "Quantitative wave-particle duality and nonerasing quantum erasure," *Phys. Rev. A*, vol. 60, pp. 4285–4290, Dec 1999.

[133] E. F. Erickson and R. M. Brown, "Calculation of fringe visibility in a laser-illuminated interferometer," *J. Opt. Soc. Am.*, vol. 57, pp. 367–371, Mar 1967.

[134] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?," *Physical review*, vol. 47, no. 10, p. 777, 1935.

[135] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, "Violation of bell's inequality under strict einstein locality conditions," *Phys. Rev. Lett.*, vol. 81, pp. 5039–5043, Dec 1998.

[136] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland, "Experimental violation of a bell's inequality with efficient detection," *Nature*, vol. 409, pp. 791–794, Feb 2001.

[137] B. S. Cirel'son, "Quantum generalizations of bell's inequality," *Letters in Mathematical Physics*, vol. 4, no. 2, pp. 93–100, 1980.

[138] R. Colbeck, "Quantum and relativistic protocols for secure multi-party computation," *PhD dissertation, Univ. Cambridge*, 2009.

[139] R. Gallego, N. Brunner, C. Hadley, and A. Acín, "Device-independent tests of classical and quantum dimensions," *Phys. Rev. Lett.*, vol. 105, p. 230501, Nov 2010.

[140] W. Mauerer, C. Portmann, and V. B. Scholz, "A modular framework for randomness extraction based on trevisan's construction," *arXiv preprint arXiv:1212.0520*, 2012.

[141] S. P. Vadhan *et al.*, "Pseudorandomness," *Foundations and Trends® in Theoretical Computer Science*, vol. 7, no. 1–3, pp. 1–336, 2012.

[142] P. Bromiley, N. Thacker, and E. Bouhova-Thacker, “Shannon entropy, renyi entropy, and information,” *Statistics and Inf. Series (2004-004)*, vol. 9, pp. 10–42, 2004.

[143] S. P. Vadhan, “Pseudorandomness,” *Foundations and Trends® in Theoretical Computer Science*, vol. 7, no. 1–3, pp. 1–336, 2012.

[144] R. Raz, O. Reingold, and S. Vadhan, “Extracting all the randomness and reducing the error in trevisan’s extractors,” *Journal of Computer and System Sciences*, vol. 65, no. 1, pp. 97–128, 2002.

[145] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, “Leftover hashing against quantum side information,” *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5524–5535, 2011.

[146] L. Trevisan, “Extractors and pseudorandom generators,” *J. ACM*, vol. 48, pp. 860–879, July 2001.

[147] X. Ma, F. Xu, H. Xu, X. Tan, B. Qi, and H.-K. Lo, “Postprocessing for quantum random number generators: entropy evaluation and randomness extraction,” *ArXiv*, 2012.

[148] X. Ma, F. Xu, H. Xu, X. Tan, B. Qi, and H.-K. Lo, “Postprocessing for quantum random-number generators: Entropy evaluation and randomness extraction,” *Phys. Rev. A*, vol. 87, p. 062327, Jun 2013.

[149] J. Zaman and R. Ghosh, “A review study of nist statistical test suite: Development of an indigenous computer package,” *arXiv preprint arXiv:1208.5740*, 2012.

[150] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, and E. Barker, “A statistical test suite for random and pseudorandom number generators for cryptographic applications,” 2001.

[151] A. Stefanov, N. Gisin, O. Guinnard, L. Guinnard, and H. Zbinden, “Optical quantum random number generator,” *Journal of Modern Optics*, vol. 47, no. 4, pp. 595–598, 2000.

[152] Trifonov and Vig, “Quantum noise random number generator,” 10 2007.

[153] Y. Shi, B. Chng, and C. Kurtsiefer, “Random numbers from vacuum fluctuations,” *Applied Physics Letters*, vol. 109, no. 4, p. 041101, 2016.

[154] T. Steinle, J. N. Greiner, J. Wrachtrup, H. Giessen, and I. Gerhardt, “Unbiased all-optical random-number generator,” *Phys. Rev. X*, vol. 7, p. 041050, 2017.

[155] F. Xu, B. Qi, X. Ma, H. Xu, H. Zheng, and H.-K. Lo, "Ultrafast quantum random number generation based on quantum phase fluctuations," *Opt. Express*, vol. 20, pp. 12366–12377, May 2012.

[156] C. Abellán, W. Amaya, M. Jofre, M. Curty, A. Acín, J. Capmany, V. Pruneri, and M. W. Mitchell, "Ultra-fast quantum randomness generation by accelerated phase diffusion in a pulsed laser diode," *Opt. Express*, vol. 22, pp. 1645–1654, Jan 2014.

[157] X. Chen, J. N. Greiner, J. Wrachtrup, and I. Gerhardt, "single-photon randomness based on a defect center in diamond," *Scientific Reports*, vol. 9, no. 1, p. 18474, 2019.

[158] L. Oberreiter and I. Gerhardt, "Light on a beam splitter: More randomness with single-photons," *Laser & Photonics Reviews*, vol. 10, pp. 108–115, Jan. 2016.

[159] H. Paul, "Photon antibunching," *Rev. Mod. Phys.*, vol. 54, pp. 1061–1102, Oct 1982.

[160] R. Renner and S. Wolf, "Simple and tight bounds for information reconciliation and privacy amplification," in *Advances in Cryptology - ASIACRYPT 2005* (B. Roy, ed.), (Berlin, Heidelberg), pp. 199–216, Springer Berlin Heidelberg, 2005.

[161] R. Loudon, *The quantum theory of light.* OUP Oxford, 2000.

[162] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light.* Elsevier, 2013.

[163] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, 1963.

[164] F. Henkel, M. Krug, J. Hofmann, W. Rosenfeld, M. Weber, and H. Weinfurter, "Highly efficient state-selective submicrosecond photoionization detection of single atoms," *Phys. Rev. Lett.*, vol. 105, p. 253001, Dec 2010.

[165] R. Colbeck, *Quantum And Relativistic Protocols For Secure Multi-Party Computation.* PhD thesis, University of Cambridge, 2009.

[166] N. Brunner, S. Pironio, A. Acin, N. Gisin, A. A. Méthot, and V. Scarani, "Testing the dimension of hilbert spaces," *Phys. Rev. Lett.*, vol. 100, p. 210503, May 2008.

[167] X. Li, A. B. Cohen, T. E. Murphy, and R. Roy, "Scalable parallel physical random number generator based on a superluminescent led," *Opt. Lett.*, vol. 36, pp. 1020–1022, Mar 2011.

[168] M. Pawłowski and N. Brunner, “Semi-device-independent security of one-way quantum key distribution,” *Phys. Rev. A*, vol. 84, p. 010302, Jul 2011.

[169] H.-W. Li, P. Mironowicz, M. Pawłowski, Z.-Q. Yin, Y.-C. Wu, S. Wang, W. Chen, H.-G. Hu, G.-C. Guo, and Z.-F. Han, “Relationship between semi- and fully-device-independent protocols,” *Phys. Rev. A*, vol. 87, p. 020302, Feb 2013.

[170] C. H. Bennett, D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and W. K. Wootters, “Remote state preparation,” *Phys. Rev. Lett.*, vol. 87, p. 077902, Jul 2001.

[171] C. H. Bennett, P. Hayden, D. W. Leung, P. W. Shor, and A. Winter, “Remote preparation of quantum states,” *IEEE Transactions on Information Theory*, vol. 51, pp. 56–74, Jan 2005.

[172] H.-K. Lo, “Classical-communication cost in distributed quantum-information processing: A generalization of quantum-communication complexity,” *Phys. Rev. A*, vol. 62, p. 012313, Jun 2000.

[173] A. Acín, S. Massar, and S. Pironio, “Randomness versus nonlocality and entanglement,” *Phys. Rev. Lett.*, vol. 108, p. 100402, Mar 2012.

[174] H.-W. Li, M. Pawłowski, Z.-Q. Yin, G.-C. Guo, and Z.-F. Han, “Semi-device-independent randomness certification using $n \rightarrow 1$ quantum random access codes,” *Phys. Rev. A*, vol. 85, p. 052308, May 2012.

[175] N. Nisan and A. Ta-Shma, “Extracting randomness: A survey and new constructions,” *Journal of Computer and System Sciences*, vol. 58, no. 1, pp. 148 – 173, 1999.

[176] P. Neumann, N. Mizuochi, F. Rempp, P. Hemmer, H. Watanabe, S. Yamasaki, V. Jacques, T. Gaebel, F. Jelezko, and J. Wrachtrup, “Multipartite entanglement among single spins in diamond,” *Science*, vol. 320, no. 5881, pp. 1326–1329, 2008.

[177] I. Vlasov, V. G. Ralchenko, A. Khomich, S. Nistor, D. Shoemaker, and R. Khmelnitskii, “Relative abundance of single and vacancy-bonded substitutional nitrogen in cvd diamond,” *physica status solidi (a)*, vol. 181, no. 1, pp. 83–90, 2000.

[178] J. Meijer, B. Burchard, M. Domhan, C. Wittmann, T. Gaebel, I. Popa, F. Jelezko, and J. Wrachtrup, “Generation of single color centers by focused nitrogen implantation,” *Applied Physics Letters*, vol. 87, no. 26, p. 261909, 2005.

[179] G. Balasubramanian, P. Neumann, D. Twitchen, M. Markham, R. Kolesov, N. Mizuochi, J. Isoya, J. Achard, J. Beck, J. Tissler, *et al.*, “Ultralong spin coherence time in isotopically engineered diamond,” *Nature materials*, vol. 8, no. 5, pp. 383–387, 2009.

[180] L. P. McGuinness, Y. Yan, A. Stacey, D. A. Simpson, L. T. Hall, D. Maclaurin, S. Prawer, P. Mulvaney, J. Wrachtrup, F. Caruso, *et al.*, “Quantum measurement and orientation tracking of fluorescent nanodiamonds inside living cells,” *Nature nanotechnology*, vol. 6, no. 6, pp. 358–363, 2011.

[181] T. Häberle, D. Schmid-Lorch, K. Karrai, F. Reinhard, and J. Wrachtrup, “High-dynamic-range imaging of nanoscale magnetic fields using optimal control of a single qubit,” *Phys. Rev. Lett.*, vol. 111, p. 170801, Oct 2013.

[182] C. E. Bradley, J. Randall, M. H. Abobeih, R. C. Berrevoets, M. J. Degen, M. A. Bakker, M. Markham, D. J. Twitchen, and T. H. Taminiau, “A ten-qubit solid-state spin register with quantum memory up to one minute,” *Phys. Rev. X*, vol. 9, p. 031045, Sep 2019.

[183] P. Neumann, J. Beck, M. Steiner, F. Rempp, H. Fedder, P. R. Hemmer, J. Wrachtrup, and F. Jelezko, “Single-shot readout of a single nuclear spin,” *Science*, vol. 329, no. 5991, pp. 542–544, 2010.

[184] A. Batalov, C. Zierl, T. Gaebel, P. Neumann, I.-Y. Chan, G. Balasubramanian, P. R. Hemmer, F. Jelezko, and J. Wrachtrup, “Temporal coherence of photons emitted by single nitrogen-vacancy defect centers in diamond using optical rabi-oscillations,” *Phys. Rev. Lett.*, vol. 100, p. 077401, Feb 2008.

[185] A. Uchida, K. Amano, M. Inoue, K. Hirano, S. Naito, H. Someya, I. Oowada, T. Kurashige, M. Shiki, S. Yoshimori, K. Yoshimura, and P. Davis, “Fast physical random bit generation with chaotic semiconductor lasers,” *Nature Photonics*, vol. 2, pp. 728–732, Dec. 2008.

[186] C. R. S. Williams, J. C. Salevan, X. Li, R. Roy, and T. E. Murphy, “Fast physical random number generator using amplified spontaneous emission,” *Opt. Express*, vol. 18, pp. 23584–23597, Nov 2010.

[187] T. Gehring, C. Lupo, A. Kordts, D. S. Nikolic, N. Jain, T. Rydberg, T. B. Pedersen, S. Pirandola, and U. L. Andersen, “Ultra-fast real-time quantum random number generator with correlated measurement outcomes and rigorous security certification,” *arXiv preprint arXiv:1812.05377*, 2018.

[188] B. Bai, J. Huang, G.-R. Qiao, Y.-Q. Nie, W. Tang, T. Chu, J. Zhang, and J.-W. Pan, “18.8 gbps real-time quantum random number generator with a photonic integrated chip,” *Applied Physics Letters*, vol. 118, no. 26, p. 264001, 2021.

[189] Y. Guo, Q. Cai, P. Li, Z. Jia, B. Xu, Q. Zhang, Y. Zhang, R. Zhang, Z. Gao, K. A. Shore, and Y. Wang, “40 gb/s quantum random number generation based on optically sampled amplified spontaneous emission,” *APL Photonics*, vol. 6, no. 6, p. 066105, 2021.

[190] D. A. Hopper, H. J. Shulevitz, and L. C. Bassett, “Spin readout techniques of the nitrogen-vacancy center in diamond,” *Micromachines*, vol. 9, no. 9, 2018.

[191] B. D’Anjou and W. A. Coish, “Optimal post-processing for a generic single-shot qubit readout,” *Phys. Rev. A*, vol. 89, p. 012313, Jan 2014.

Acknowledgment

Throughout the writing of this dissertation, I have received a great deal of support and assistance. This Ph.D. thesis would not have been possible without the support and assistance I got from my colleagues, my friends, and my family. All I would like to say is a special thank you to you all.

- Prof. Dr. Jörg Wrachtrup, first, thank you very much for giving me the opportunity to conduct my Ph.D. research in your group and providing me with all the world-class research facilities. And thank you for your continuous support in both my academic and non-academic life.

- Prof. Dr. Ilja Gerhardt, thank you for your guidance and support in conducting all the randomness generation experiments and writing the papers. And thank you for your detailed suggestions to improve this thesis.

- Prof. Dr. Stefanie Barz, thank you for being the Second examiner of my thesis.

- Prof. Dr. Thomas Speck, thank you for being the Chairperson of my Examination.

- Oliver, Santo, Vlad, Dr. Florian Kaiser, Dr. Jianpei Geng, Dr. Durga Dasari. Thank you all for providing the initial feedback on this thesis.

- All the members in PI3, thank you all for being there in the past four years and being supportive whenever I needed help.

- Thank my colleagues in Q.ANT, especially Dr. Tobias Wintermantel for providing me valuable comments and suggestions to my thesis.

- Especially thank Santo and Oliver for being great office-mates and supporting me with daily trivial and non-trivial matters. It is a very good memory to spend with you two for the last few years in the same office.

- My parents, my brother, and my sisters thank you all for supporting me unconditionally whenever and wherever I am.

- Zhen, my wife, thank you for being with me, your unlimited love for me is the biggest motivation for me to do a Master's study and eventually a Ph.D. study abroad. Without your support, I cannot imagine where I could be now.

Curriculum Vitae

A brief academic curriculum vitae of myself is included here, a more detailed CV can be found at `https://www.linkedin.com/in/xing-chen-b6626070b/`

Dr. rer. nat. in Physics

- University of Stuttgart, May 2018 - Jan 2022
- Research field: Quantum information theory, quantum randomness
- Thesis: Quantum Randomness Certified By Different Quantum Phenomena

Master of Science in Theoretical Physics

- University of Chinese Academy of Sciences, Sept 2014 - July 2017
- Research field: Quantum information theory, quantum entanglement
- Thesis: Quantum entanglement and nonlocal game

Bachelor of Science in Mathematics

- Beijing Institute of Technology, Sept 2010 - July 2014
- Major: Mathematics and applied mathematics
- Thesis: Fast Fourier Transform Algorithm for Two Types of Partial Differential Equations on Rectangular Domains