

A new approach to optimize the transferable
anisotropic Mie force field (TAMie)
for mixtures

A thesis accepted by the Faculty of
Energy-, Process- and Bio-Engineering of the University of Stuttgart
in partial fulfilment of the requirements for the degree of
Doctor of Engineering Sciences (Dr.-Ing.)

by

Dominik Weidler, M.Sc.

born in Biberach an der Riß

Committee chair: Prof. Dr.-Ing. Joachim Groß

Committee member: Prof. Richard J. Sadus

Date of defence: 20.12.2021

Institute of Thermodynamics and
Thermal Process Engineering
University of Stuttgart
2021

“The highest forms of understanding we can achieve are laughter and human compassion.”
(Richard Feynman)

This book is dedicated to my family, my friends, my colleagues and all the many other people who enrich my life and stand by me in good and in desperate times. Thank you for supporting me.

List of Abbreviations

- 2D** Two-dimensional
- 3D** Three-dimensional
- AAD** Absolute Average Deviation
- AD** Average Deviation
- AMBER** Assisted Model Building with Energy Refinement
- AUA4** A variation of the AUA force field
- AUA** Anisotropic United Atom
- BPS** Branch-Point Sampling
- CASSANDRA** An Open Source Monte Carlo Package for Molecular Simulation
- CBMC** Configurational Bias Monte Carlo
- CB** Configurational Bias
- CHARMM** Chemistry at Harvard Macromolecular Mechanics
- CPU** Central Processing Unit
- DDBST** Dortmund Data Base
- DDB** Dortmund Data Base
- DIADEM** A front-end for the DIPPR database
- DIPPR** Desing Institute for Physical Properties
- DNA** Deoxyribonucleic acid
- EOS** Equation of state
- GCMC** Grand Canonical Monte Carlo (simulations)
- iTAMie** individualized TAMie
- KDB** Korean Thermophysical Properties data base
- LJ** Lennard-Jones
- MBAR** Multistate Bennett Acceptance RatioMBAR

MC Monte Carlo

MD Molecular Dynamics

MP2 Møller-Plesset perturbation theory of second order

MPI Message Passing Interface

NERD Force field named after S.K. Nath, F.A. Escobedo, J.J. de Pablo

OPLS-AA Optimized Potentials for Liquid Simulations All Atomistic

OPLS Optimized Potentials for Liquid Simulations

PC-SAFT Perturbed Chain - Statistical Associating Fluid Theory

RMSD Root-Mean-Square Deviation

RMS Sum of squared deviations

SAFE Self-Adapting Fixed-Endpoint

SAFT Statistical Associating Fluid Theory

TAMie Transferable anisotropic Mie (potential)

TraPPE-EH Transferable Potentials for Phase Equilibria with Explicit Hydrogens

TraPPE Transferable Potentials for Phase Equilibria

TURBOMOLE An application for quantum chemical calculations

UA United Atom

VLE Vapor Liquid Equilibria

VR-SAFT Statistical Associating Fluid Theory for chain molecules with attractive potentials of Variable Range

VR Variable Range

List of Symbols

Symbol	Description
ψ, ψ_A	Correction parameter of the iTAMie force field
E	(internal) Energy
T	Temperature
p	Pressure or probability (distribution)
V	Volume
N	Number of molecules (in a system)
N_i	Number of molecules of a substance i
Δt	Integration time step
\mathbf{r}^N	Vector of all positions of all interaction sites
\mathbf{p}^N	Vector of all momenta of all interaction sites
D	Diffusion coefficient
ϵ_i	Energy of a molecule i
μ	Chemical potential
Π	(projection of a) probability distribution
$\Delta_k N$	Size of a window k in a simulation
H	Histogram
$H(N, E)$	Two-dimensional histogram of number of molecules and energy
π	probability (of being in a microstate)
β	thermodynamik beta or inverse temperature $\beta = 1/(k_B T)$
k_B	Boltzmann constant
Λ	The de Broglie wavelength $\Lambda(T)$
Ξ	Grand partition function
$a_{x \rightarrow y}$	Biased acceptance probability from x to y
ω	Bias function $\omega(N)$
C	Collection matrix of the transition matrix approach
P	Biased transition probability
δ	Vector $\delta = \{-1, 0, 1\}$
ΔE	Width of an energy bin
Ω	Microcanonical partition function
C_i	Constant C_i being the fraction of two grand partition functions
$\bar{\beta}$	Averaged thermodynamic beta
I	Sum over all energies and molecule numbers
$\langle A \rangle$	Ensemble average of A
p^{sat}	Vapor pressure
Γ	Part of the scaling factor $\exp(\Gamma)$
$\exp(\Gamma)$	Scaling factor to combine histograms of different simulations
ρ	Density
ρ^*	Dimensionless density $\rho^* = \rho/\rho_{\text{max}}$

\wedge	Mathematical AND
W_n	Rosenbluth weight
A	Helmholtz free energy
\hat{a}	Dimensionless Helmholtz free energy $\hat{a} = A/Nk_B T$
Z	Compressibility factor
σ	Size parameter of a group contribution method
ε	Energy parameter of a group contribution method
Ω^{sim}	Computed properties of interest
Ω^{exp}	Properties of interest for experiments
\mathbf{p}	Vector of force field parameters
$f(\mathbf{p})$	Objective function to be minimized
f_{aux}	Auxiliary objective function to be minimized
κ_{ij}	Group-wise cross interaction parameter (groups i and j)
$\kappa_{\alpha\beta}$	Substance-wise cross interaction parameter (substances α and β)
n_{ij}	Repulsive exponent of a Mie potential between interaction sites i and j
x_α, x_β	Amount of substance fraction for substances α and β
$\varphi_{\alpha\beta}$	Correction parameter
T_R	Reduced temperature $T_R = T/T_{\text{exp}}^{\text{crit}}$
$T_{\text{exp}}^{\text{crit}}$	Critical temperature from experimental data
m_{ij}	Dispersive exponent of a Mie potential between interaction sites i and j
r_{ij}	Distance between two interaction sites
u	Potential
u_{ij}^{inter}	Intermolecular potential between two interaction sites
c_{ij}	A constant to ensure the minimum of the Mie potential to be at a value of $-\varepsilon_{ij}$
e	Charge of an electron
r_c, R_c	Cutoff radius
θ	Bending angle between three consecutive interaction sites
θ_{eq}	Equilibrium angle of a bending potential
k_0	Force constant of a bending potential
ϕ	Actual angle of a dihedral angle
c_0, c_1, c_2, c_3	Force constants of torsional potentials
\in	Is member of
α	Damping parameter of the Ewald summation
k_{max}	Maximum number of k-vectors of the Ewald summation
f_{chair}	Probability of finding a chair confirmation of a cyclic molecule
f_{boat}	Probability of finding a boat confirmation of a cyclic molecule
q_x	Point charge of an interaction site
ε_0	Vacuum permittivity
T_c	Critical temperature
\forall	For all
$g(r)$	Radial pair distribution function
\tilde{r}	Reduced radial distance
\tilde{u}	Reduced potential function
\hat{x}	Parameters of the PC-SAFT eos, where $x \in \{m, \sigma, \varepsilon, g\}$

Subscripts	Description
<i>i</i>	Species i
<i>k</i>	A specific condition in a GCMC simulation
<i>j</i>	Generic index
ov	At overlapping conditions
1,2	Substances 1 and 2
<i>f</i>	Generic index
<i>o</i> → <i>n</i>	Old to new
<i>c</i>	Critical
cor	Correlation
Superscripts	Description
\bar{x}	Vector of all substances of a variable x or average of a variable
<i>I</i>	Phase I
<i>II</i>	Phase II
coex	Coexisting phases (e.g. vapor-liquid equilibrium)
ex	Excess property
bias	Biased variable
min	Minimum value
max	Maximum value
L,l,liq	Liquid
V,v,vap	Vapor
sep	Separation point (Location where the probability distribution has its minimum between two phases)
sat	Saturation
W	Currently considered window
inter	Intermolecular
intra	Intramolecular
hc	Hard chains
disp	Dispersive
assoc	Associating
polar	Polar
res	Residual
exp	Experiment
sim	Simulation
ind	Individualized
fs	Überlappende Kugeln (fused spheres)

Contents

List of Abbreviations	1
List of Symbols	3
List of Figures	9
List of Tables	11
1. Introduction	16
2. Theory and Methods	19
2.1. Introduction to molecular simulations	19
2.1.1. Monte Carlo vs. Molecular Dynamics	19
2.1.2. Different ensembles and their applications	21
2.2. Grand Canonical Monte Carlo simulations to determine phase equilibria	23
2.2.1. Transition matrix Monte Carlo and multiple histograms	24
2.2.2. Histogram reweighting for pure substances	29
2.2.3. Histogram Reweighting for binary mixtures	31
2.2.4. Configurational bias Monte Carlo (CBMC)	34
2.2.5. A fragment library for branched and cyclic molecules	35
2.3. The PC-SAFT equation of state and its application to force field optimization	38
2.3.1. The perturbed chain statistical associating fluid theory (PC-SAFT) equation of state	38
2.3.2. Generating conditions for GCMC simulations with PC-SAFT	39
2.3.3. Optimization of Mie parameters with PC-SAFT	39
2.3.4. Determining cross interaction parameters with PC-SAFT	40
3. TAMie for aldehydes, ketones and small cyclic alkanes	43
3.1. Introduction	43
3.2. Force field development	46
3.2.1. Intermolecular energy	46
3.2.2. Intramolecular energy	47
3.2.3. Optimization of force field parameters	49
3.2.4. Simulation details	49
3.3. Results	51
3.3.1. Small cyclic alkanes	51
3.3.2. Ketones	52
3.3.3. Aldehydes	55
3.4. Transferability of parameters	57
4. Individualized force fields based on the transferable anisotropic Mie potential	60
4.1. Introduction	61

4.2. Methodology	63
4.2.1. The TAMie force field	63
4.2.2. The iTAMie force field	64
4.2.3. Simulation Details	66
4.3. Results	67
4.3.1. Assessment of the scaling method	67
4.3.2. Alkanes and olefins	69
4.3.3. Ethers	69
4.3.4. Ketones	71
4.3.5. Application to a binary mixture	71
4.4. Conclusion	73
5. Phase equilibria of binary mixtures with alkanes, ketones, and esters based on TAMie	74
5.1. Introduction	75
5.2. Force field development	76
5.2.1. <i>Ab initio</i> calculations of intramolecular bending and torsional potentials	76
5.2.2. Optimization of intermolecular force field parameters for esters	77
5.3. Simulation details	81
5.3.1. Pure substances	81
5.3.2. Binary mixtures	82
5.3.3. Cross-energy parameters κ_{ij} for binary mixtures	83
5.4. Results	83
5.4.1. Intramolecular potential	83
5.4.2. Pure esters	85
5.4.3. Individualized pure esters	89
5.4.4. Binary mixtures	91
5.5. Conclusion	93
6. Conclusion and outlook	97
A. CPU time usage	100
B. TAMie for benzene	102
B.1. Dynamic properties of benzene	104
C. Expressing the PC-SAFT model in terms of force field pure parameters	107
D. Influence of the energy bin width on results at low temperatures	110
D.1. Background to the energy bin width	110
D.2. Bin sizes used in this thesis	111
E. Supporting information for chapter 3	113
E.1. Background to reported results	113
E.1.1. Errors	113
E.1.2. Enthalpies of vaporization	113
E.1.3. Critical points	114

E.2. Results for cyclopentane and cyclohexane	114
E.3. Results for ketones	115
E.4. Results for aldehydes	118
E.5. Confirming simulation results with other simulation software	121
F. Supporting information for chapter 4	122
F.1. Alternative scaling schemes	122
F.2. Simulation Details	124
F.3. Correlation function parameters for the newly fit vapor pressure correlation of ethyl-propyl-ether	124
F.4. Enthalpies of vaporization	126
F.5. Graphical results	126
G. Supporting information for chapter 5	141
G.1. Background to reported results	141
G.1.1. Enthalpies of vaporization	141
G.1.2. Critical points	141
G.2. Simulation Details	141
G.3. Results for formates	143
G.4. Results for acetates	144
Bibliography	146

List of Figures

2.1. Unbiased probability distribution	24
2.2. Schematic depiction of the probability distribution $\Pi(N; \mu, T)$ and the division into several windows k with window size $\Delta_k N$	26
2.3. Biased probability distribution	27
2.4. Probability distribution at different conditions	28
2.5. Different simulation conditions along the phase envelope for simulations in grand canonical ensemble [1].	29
2.6. Cases to determine Γ	33
2.7. Illustration of a growing step with the fragment library	37
2.8. Crank-shaft moves for a cyclic molecule	37
3.1. Vapor pressure of cyclopentane and cyclohexane	53
3.2. VLE coexisting densities for cyclic alkanes	53
3.3. Relative errors of cyclopentane and cyclohexane	54
3.4. AADs of acetone for different partial charges	55
3.5. AADs of aldehydes for different partial charges	57
3.6. Phase diagrams of aldehydes	58
3.7. Vapor pressures of aldehydes	58
3.8. Relative deviations of ketones	59
4.1. Relative errors of TAMie for ethers	63
4.2. AAD for p^{sat} of TAMie and iTAMie	68
4.3. AD for p^{sat} of TAMie and iTAMie	68
4.4. Deviations of TAMie and iTAMie for two ethers	70
4.5. Relative deviations of ethyl-propyl-ether to DIPPR and our correlation function	70
4.6. Benefit of iTAMie for p^{sat} of ketones	71
4.7. Binary mixture of butanone and 1-hexene at $T = 333$ K	73
5.1. Torsional energies for the O=CH–O–CH ₂ dihedral of ethyl formate	78
5.2. Torsional energies for the CH _x –O–CH ₂ –CH ₃ dihedral angle in esters	78
5.3. Torsional potentials for ethyl acetate	79
5.4. RMS deviations for formates	86
5.5. Comparison of vapor pressures for formates	87
5.6. Relative deviations of the vapor pressure for formates	87
5.7. Relative deviations of calculated liquid densities for formates	88
5.8. RMS deviations for acetates	88
5.9. Relative deviations of the vapor pressure for acetates	90
5.10. Relative deviations of the liquid densities for acetates	90
5.11. Comparison of VLE data in a T - ρ -projection for acetates	91

5.12. Phase diagram of ethyl acetate with heptane at $T = 343.15$ K	92
5.13. Vapor liquid equilibrium of butanone with n -hexane at $T = 333.15$ K	94
5.14. Vapor liquid equilibrium of butanone with n -heptane at $T = 323.15$ K . . .	95
5.15. Vapor liquid equilibrium of 2-pentanone with n -heptane at $T = 363.15$ K . .	95
5.16. Excess enthalpies h^E for the mixture of butanone and hexane at $T = 298.15$ K	96
A.1. CPU time usage for MPI implementation without optimization	100
A.2. CPU time usage for MPI implementation with optimization	101
B.1. Structure of a UA benzene molecule	102
B.2. Relative deviations of the saturation pressure of benzene	103
B.3. Relative deviations of the liquid densities of benzene	103
B.4. Diffusion coefficients of benzene	104
B.5. Shear viscosity of benzene	105
B.6. Thermal conductivity of benzene	105
D.1. Relative deviations of p^{sat} for ethylacetate at different energy bin widths .	111
D.2. Relative deviations of p^{sat} for ethylacetate at different energy bin widths .	112
E.1. Phase diagram of simulated ketones	117
E.2. Ketones: Enthalpies of vaporization	117
E.3. Vapor pressures of ketones	118
E.4. Vapor pressures of aldehydes	120
E.5. Aldehydes: Enthalpies of vaporization	120
F.1. Scaling behavior of ketones when scaling the Mie energy parameter	123
F.2. Scaling behavior of ketones when scaling the Mie size parameters	123
F.3. AAD combined for all individualized substances	127
F.4. AAD of the vapor pressure for all individualized substances	127
F.5. AAD of the liquid density for all individualized substances	128
F.6. AAD of the liquid density for all individualized substances	128
F.7. AD of the vapor pressure for all individualized substances	129
F.8. AD of the liquid density for all individualized substances	129

List of Tables

2.1. Classical ensemble types	22
2.2. Number of CB steps for different dimensionless densities $\rho^* = \rho/\rho_{\max}$	35
3.1. Pair potential parameters of UA-groups: energy parameter ε , size parameter σ , repulsive exponent n , and point charge q for the TAMie force field	46
3.2. Bond lengths used in the proposed TAMie force field taken from literature.	48
3.3. Bending angles and constants for the proposed TAMie force field taken from literature.	48
3.4. Torsional potential constants taken from literature.	48
3.5. AADs of TAMie for ketones and aldehydes	56
3.6. AADs of TAMie-predictions for ketones and aldehydes	59
4.1. Pair potential parameters of UA-groups: energy parameter ε , size parameter σ , repulsive exponent n , and point charge q for the TAMie force field	65
4.2. Bond lengths used in the proposed TAMie force field taken from literature.	65
4.3. Bending angles and constants for the proposed TAMie force field taken from literature.	65
4.4. Torsional potential constants taken from literature.	65
4.5. Comparison of different methods for individualizing the TAMie force field for two substances (butanone and di-n-pentyl-ether).	67
4.6. Optimization parameters ψ_A and AAD of TAMie and iTAMie	72
5.1. Bond lengths of the TAMie force field for esters.	84
5.2. Angle bending parameters of the TAMie force field for esters.	84
5.3. Torsional potential parameters of the TAMie force field for esters.	84
5.4. van der Waals (Mie) parameters and point charges of the TAMie force field for esters.	89
5.5. Comparison of the combined absolute average deviations for the TAMie and iTAMie force field.	92
B.1. van der Waals (Mie) parameters, bond length l and bending angle θ of the TAMie force field for benzene.	103
D.1. Tabulated energy bin widths	112
E.2. Critical properties of cyclic alkanes	114
E.1. Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization for cyclic alkanes	115
E.3. Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization for ketones	115
E.4. Critical properties of simulated ketones	116

E.5.	Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization for aldehydes	118
E.6.	Critical properties of simulated aldehydes	119
E.7.	Consistency check using GEMC with CASSANDRA for 2-pentanone. 10 Mio equilibration steps + 10 Mio production steps in Gibbs Ensemble. Errors are calculated using the reblocking method of Flyvbjerg and Petersen[2].	121
F.1.	simulation details	125
F.2.	Number of CB steps for different dimensionless densities	126
F.3.	Parameters for the Clausius-Clapeyron like fit of the vapor pressure for ethyl-propyl-ether	126
F.4.	Optimization parameters ψ_A and AAD of TAMie and iTAMie compared to DIPPR correlations for different substances. Ethyl-propyl-ether* has been adjusted to quasi-experimental data of a newly fit correlation function.	130
F.5.	Alkanes: Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization. The last line for each substance gives the estimated critical properties.	131
F.6.	Olefins: Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization. The last line for each substance gives the estimated critical properties.	135
F.7.	Ether and Ketones: Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization. The last line for each substance gives the critical properties.	138
G.1.	simulation details	142
G.2.	Number of CB steps for different dimensionless densities	142
G.3.	Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization for formates	143
G.4.	Critical properties of simulated formates	143
G.5.	Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization for acetates	144
G.6.	Critical properties of simulated acetates	144

Journal publications

This thesis led to the following publications:

- Chapter 3: D. Weidler and J. Gross. Transferable anisotropic united-atom force field based on the mie potential for phase equilibria: aldehydes, ketones and small cyclic alkanes. *Ind. Eng. Chem. Res.*, 55(46):12123–12132, 2016
- Chapter 4: D. Weidler and J. Gross. Individualized force fields for alkanes, olefins, ethers and ketones based on the transferable anisotropic mie potential. *Fluid Phase Equilib.*, 470:102–108, 2018
- Chapter 5: D. Weidler and J. Gross. Phase equilibria of binary mixtures with alkanes, ketones, and esters based on the transferable anisotropic mie force field. *Fluid Phase Equilib.*, 490:123–132, 2019

The chapters 3 to 5 present literal quotes of the published work. Any addition with respect to the published work is marked. Any deletion is indicated with square brackets as ,[...]'. Cross-references between chapters of this thesis, which are added to the published version of the text to increase readability, are marked by square brackets. The SupportingInformation is presented in the Appendix of this thesis.

Kurzfassung

In dieser Arbeit wird die Entwicklung eines molekularen Kraftfeldes dargestellt, das einen Anwendungsschwerpunkt auf der Berechnung von Phasengleichgewichten zulässt. Aufbauend auf dem „Transferable Anisotropic Mie (TAMie)“ Kraftfeld von Hemmen et al. [1, 6, 7] wird der Parametersatz des Kraftfeldes auf kleine zyklische Moleküle sowie polare Stoffgruppen wie Ester und Ketone ausgeweitet. Es handelt sich um ein klassisches, atomistisches Kraftfeld, bei dem jedoch Wasserstoffatome häufig effektiv mit benachbarten größeren Atomen zusammen berücksichtigt werden. Die Kraftfeldparameter sind dabei übertragbar, d.h. sie können innerhalb einer Stoffgruppe für alle Substanzen verwendet werden. Obwohl die erzielten Ergebnisse der Phasengleichgewichte sehr gut sind, stößt das übertragbare Modell mit einfachen Punktladungen dabei an Grenzen. Dies zeigt sich in Abweichungen des Sättigungsdampfdruckes aus den Simulationen gegenüber experimentellen Daten. Eine möglichst genaue Beschreibung des Dampfdruckes ist allerdings erstrebenswert, um Mischungseigenschaften mit guter Übereinstimmung zu experimentellen Daten vorhersagen zu können. Um den übertragbaren Charakter des Kraftfeldes nicht zu zerstören und gleichzeitig die Genauigkeit des Dampfdruckes für einzelne Stoffe zu erhöhen wird das individualisierte TAMie Kraftfeld eingeführt. Dabei werden mit Hilfe eines Korrekturparameters ψ alle energetischen Wechselwirkungen eines Reinstoffes skaliert, um für experimentell gut vermessene Stoffe die Genauigkeit zu erhöhen. Es wird gezeigt, dass dieses Konzept zu deutlich verbesserten Korrelationen und Vorhersagen von Mischungseigenschaften führt. Anhand verschiedener binärer Mischungen wird auch die Übertragbarkeit von Kreuzwechselwirkungsparametern, die die van-der-Waalschen Wechselwirkungen zwischen zwei Reinstoffen korrigieren, gezeigt. Es werden weitere Untersuchungen und Experimente zur Validierung empfohlen.

Abstract

In this thesis the development of a molecular force field is presented, which allows an application focus on the calculation of phase equilibria. Based on the "Transferable Anisotropic Mie (TAMie)" force field by Hemmen et al. [1, 6, 7] the parameter set of the force field is extended to small cyclic molecules and polar groups of substances such as esters and ketones. It is a classical atomistic force field, but hydrogen atoms are often effectively considered together with neighbouring larger atoms. The force field parameters are transferable, i.e. they can be used for all substances within a group of substances. Although the phase equilibrium results obtained are very good, the transferable model with simple point charges reaches some limitations. This is shown in deviations of the saturation vapor pressure from the simulations compared to experimental data. However, it is desirable to describe the vapor pressure as accurately as possible in order to be able to predict mixture properties with good agreement to experimental data. In order not to destroy the transferable character of the force field and at the same time ensure the accuracy of the vapor pressure for individual substances, the individualized TAMie force field is introduced. With the help of a correction parameter ψ all energetic interactions of a pure substance are scaled in order to increase the accuracy for experimentally well measured substances. It is shown that this concept leads to significantly improved correlations and predictions of mixture properties. Using various binary mixtures, the transferability of cross-interaction parameters that correct van der Waals interactions between two pure substances is also demonstrated. Further investigations and experiments are recommended for validation.

1. Introduction

Molecular simulations are going to play an important role in future examinations of substance properties and many other research fields. Molecular simulation refers to all computer simulations from a quantum mechanical time- and lengthscale (a few femtoseconds for the electron structure of one or a few atoms), over the microscopic time and lengthscale (several tens to thousands of molecules for the time of nano to microseconds) up to a mesoscopic time- and lengthscale, e.g. systems containing proteins (sizes up to micrometers over several seconds). Depending on the system size and the field of research, different approaches and models have been developed over the last decades.

The future applications of molecular simulations might include the design of effective and gentle drugs or may help to gain a better understanding of how metabolic processes in animals and humans work. Beside biological and medical research fields and applications, molecular simulations are also an established tool to solve engineering problems, e.g. determine substance properties of pure substances and mixtures, like phase equilibria, surface tensions or adsorption isotherms, but also dynamic properties like diffusion coefficients. These substance properties can afterwards be used to design and optimize industrial processes.

The determination of substance properties with molecular simulations is often done on the microscopic scale using classical force fields. For the rest of this thesis I will use the term “molecular simulations” for such simulations. There are numerous approaches and methods within the framework of molecular simulations with which substance properties can be determined. Excellent overviews are e.g. given in [8–18]. Chapter 2 introduces the most relevant theories and simulation techniques for this thesis.

The methods mentioned include different kinds of molecular simulations as well as the underlying molecular model, i.e. force fields. Force fields are the backbone of every molecular simulation, as they describe all interactions in and between molecules. A force field consists of ansatz functions, representing various types of bonds and interactions together with a set of corresponding parameters defining these interactions.

The level of detail of a force field depends on the anticipated applications. All-atomistic models treat every atom in a molecule as an interaction site, and give detailed insight into the microscopic structure of fluids. AMBER [19–23] (**A**ssisted **M**odel **B**uilding with **E**nergy **R**efinement) and CHARMM (**C**hemistry at **H**arvard **M**acromolecular **M**echanics)

1. Introduction

are two types of force fields which are often used in biological applications like protein folding or DNA-protein interactions. Different versions of AMBER and CHARMM exist and depending on the purpose of the investigation, the most suitable version has to be chosen. For CHARMM the use of a particular version depends e.g. on whether proteins [24–27], nucleic acids [28–31], lipids[32–34] or carbohydrates[35–38] are of special interest.

Often, force fields for biological investigations are optimized to ambient conditions or laboratory conditions, meaning pressures about 1 bar and temperatures about 25 °C. The fact that force field parameters were adjusted to properties at these limited conditions makes them inappropriate to determine phase equilibria over a wide range of pressures and temperatures.

Other force fields like the **O**ptimized **P**otentials for **L**iquid **S**imulations **a**ll **a**tomistic (OPLS-AA) force field [39] or the **T**ransferable **P**otentials for **P**hase **E**quilibria with **E**xplicit **H**ydrogens force field [40–43] (TraPPE-EH) perform much better for a wide range of pressures and temperatures. A slightly more coarse version treats hydrogen atoms around a central atom – for example the hydrogen atom of a methyl group (CH₂) – combined with the central atom as one interaction site. This approach is referred to as **U**nited-**A**tom (UA) approach. Besides the all-atom OPLS and TraPPE force fields, both models also have united-atom parameterizations (OPLS-UA [44–49] and TraPPE-UA [50–54]).

Another approach was taken by **Mueller and Jackson** who recently developed a method to directly extract force field parameters from a physically based equation of state leading to good results for thermodynamic properties [55–57]. However, it is limited to the underlying molecular model of the equation of state and is thus more coarse than atomistic force fields, leading to significant deviations in dynamic properties.

As it can plainly be seen, several different force fields have been developed and established over the past decades and each of them has advantages for a certain purpose. For example the OPLS force field, as its name implies, was developed especially to simulate liquid systems. The TraPPE force field on the other hand was developed to give accurate descriptions of phase equilibria, especially vapor liquid equilibria, in a transferable way. Transferable means, that the same chemical groups in different molecules are described with the same parameters. The goal of a transferable force field is to predict the interactions in and between molecules which were not part of the optimization procedure. This gives – with the appropriate simulation technique – access to properties that are not available as experimental data in literature or difficult to measure because the conditions are demanding for experimental work or the substances themselves are dangerous. In the future, molecular simulations might replace some experimental setups or at least will be used to support them by giving profound estimations.

1. Introduction

The work presented in this thesis follows the works of Hemmen [1, 6, 7] and coworkers during her studies at the University of Stuttgart. Many methods used in this work are described in detail in her thesis[7]. Hence, I keep the methods section brief and focus on the key methods and new developments of the **T**ransferable **A**nisotropic **M**ie (TAMie) [1, 6, 7, 58] force field and on introducing the concept of an individualized force field based on TAMie [4].

The next chapters cover the essential theoretical parts and simulation methods used to develop this force field, the application to aldehydes, ketones and small cyclic alkanes [3], the development of an individualized force field [4] and to mixtures including parameters for esters [5]. At the end of the main thesis a conclusion and outlook for further research is given. In the appendix some results are presented, that have not been published before. Appendix A gives a short comparison of CPU usage when applying a simple **M**essage **P**assing **I**nterface (MPI) scheme to multicanonical sampling. In Appendix B a TAMie force field for benzene, which works without electrostatic interactions but nonetheless gives excellent results for dynamic properties is given. The appendix also includes the detailed description of the relation between the molecular model used for molecular simulations in this thesis (TAMie) and the equation of state (PC-SAFT [59–62]) used to approximate and optimize the force field parameters. Finally, a short estimation of how the choice of the bin width for collecting histograms during simulation affects the results for low temperatures of the simulations, which was found out at the very end of the research described in this thesis.

2. Theory and Methods

In this chapter the basic principles of molecular simulations are briefly explained. This includes the description of different kinds of ensembles which are connected directly to macroscopic thermodynamic systems. Afterwards, an overview of the grand canonical Monte Carlo simulation to determine phase equilibria is given. I rationalize why Monte Carlo simulations in grand canonical ensemble are – combined with special techniques like transition matrix sampling and histogram reweighting – a good choice to determine static properties of phase equilibria for pure substances as well as for mixtures. Furthermore, I review the most important techniques to minimize statistical uncertainties in the chosen ensemble. The second part of this chapter will be an overview of the perturbed chain statistical associating fluid theory (PC-SAFT) and its connection to the united atom model of the TAMie force field. In this work PC-SAFT is used to get an initial guess for suitable control variables, at which the molecular simulations are performed and furthermore it is used for accelerating the optimization of force field parameters of TAMie. Finally, I explain the necessary basics on how to calculate dynamic properties with help of molecular dynamics simulations.

2.1. Introduction to molecular simulations

This section covers the fundamentals of molecular simulations and their applications to determine macroscopic thermodynamic properties. The concept of an ensemble will be explained, as well as the difference between Molecular Dynamics and Monte Carlo simulations. For a far broader and more detailed overview of the history and basics of molecular simulations I refer to [14–18, 63, 64].

2.1.1. Monte Carlo vs. Molecular Dynamics

The underlying principle behind all molecular simulations is that macroscopic properties can be determined by studying the microscopic behavior of a system. Every macrostate is defined by its macroscopic control variables (for example energy E , temperature T , pressure p , volume V , amount of substances N_i), depending on the thermodynamic system. Because microstates can be forced to obey the control variables, the macroscopic properties can be calculated by averaging over all microstates belonging to the chosen state

2. Theory and Methods

variables of the macroscopic system. Depending on which state variables are defined for the thermodynamic system, a different ensemble-type is used for molecular simulations. An ensemble is simply the collection of infinitely many microstates appearing according to the correct probability and all of them obey the defined control variables. For an N,V,E -ensemble or “microcanonical ensemble”, the energy of every molecule in the microscopic system can fluctuate, as long as the sum of the energy of all molecules E in the system is constant (next to constant volume V and number of molecules N). The macroscopic properties can then be determined by averaging over all microstates.

Two fundamentally different approaches are available to generate the different microstates. In Molecular Dynamics (MD) one generates an initial microstate which fulfills the constraints of the macroscopic thermodynamic system and then subsequently one solves Newtons law of motion for every interaction site in the system. Therefore, an integration over all forces acting on every interaction site is necessary and the integration time step Δt has to be sufficiently small to avoid numerical instabilities. This approach requires knowledge of all positions \mathbf{r}^N and all momenta \mathbf{p}^N for all interaction sites. To generate a new microstate the forces acting on every interaction site and the total energy of the system have to be evaluated for every time step. The average in this case is a time average.

The second approach to generate new microstates is referred to as Monte Carlo simulations. As the name implies, chance plays an important role in this approach. New microstates are generated by randomly choosing a molecule (or part of it) and also randomly applying an action to it. This action can be the change of the position of the molecule, a new internal configuration of the molecule (if the molecule is made of several interaction sites), or even remove the molecule from the simulation box or add a new one. To guarantee that any trial move leads to a microstate that is representative for the given control variables, i.e. a microstate appears according to the correct probability for this microstate, an evaluation step is required after each trial move. The evaluation step depends on the chosen ensemble. Some advantages of this approach are: “unphysical” moves can be performed, which sometimes allows for a more efficient sampling of the phase space; there is no need to evaluate the forces of the system, since moves happen randomly and it is easy to simulate at constant temperature T . The last point is one advantage of MC for the determination of phase equilibria, because one of the three requirements for thermodynamic equilibrium is the thermal equilibrium. However, constant T is only an assumption in the framework of MC and a sufficient number of displacement steps - which is not known *a priori* - is necessary and has to be verified in the end. In MD simulations on the other hand, this condition requires the application of a thermostat. The second and more important advantage of MC over MD for the determination of phase equilibria is the possibility to determine the chemical potential μ - which is the third re-

quirement for thermal equilibria - more easily, although there are also hybrid approaches [65]. The disadvantage of Monte Carlo simulations is that one does not (approximately) follow the dynamics of a system; no dynamic properties, such as diffusion coefficients or viscosity can be determined.

2.1.2. Different ensembles and their applications

Table 2.1 [66] shows classical ensemble types, their control variables, fluctuating properties and the corresponding thermodynamic system. These different ensembles are used for different kinds of thermodynamic problems.

The microcanonical ensemble with constant number of molecules N_i of all species, volume V and energy E was seldom used for the study of thermophysical properties, since both, the pressure p and the temperature T can fluctuate whereas experimental setups usually keep one of these state variables constant. However, temperatures, chemical potentials μ and diffusion coefficients D could be obtained with simulations using the microcanonical ensemble. In 1998 Lustig came up with the MD NVE approach[67] and in his approach he introduced a sampling procedure in configurational phase space in the microcanonical (NVE) ensemble to simulate any thermodynamic property. His comparisons showed equivalent accuracy of his method compared to the best available equations of state for Lennard-Jones systems. His approach was refined by Meier and Kabelec in 2006, who solved some problems reported by Lustig, i.e. the correct calculation of isentropic and isothermal compressibilities, the speed of sound and higher pressure derivatives up to second order[68].





In the canonical ensemble the number of molecules N_i of each species i , the volume V and the temperature T are kept constant. The corresponding thermodynamic system would be a closed one with fixed V and T and with the possibility to exchange energy with the surrounding environment. It is used to calculate properties like pressure or dynamic properties, if the density is known.

The isothermal-isobaric ensemble (or N,p,T-ensemble) was originally thought of to be particularly appropriate for simulating mixtures, since experimental data, e.g. excess enthalpies, are often measured at constant pressure, and the assumptions behind many theories of mixing often also require constant pressure [17]. For chemical reactions, this ensemble is useful in the same way, as these reactions are usually carried out under constant pressure [16]. However, in a N, p, T -ensemble observes two phases only for rather large systems, which makes it difficult to study phase transitions or phase equilibria [17].

The grand canonical ensemble is special in this list, since it is the only ensemble (of the ones mentioned above) where the number of molecules is not fixed. If we knew the right chemical potentials $\mu_i(T, V)$ leading to coexisting phases, say a liquid and vapor phase, the ensemble would allow to simulate the vapor and liquid phase of a substance in one

2. Theory and Methods

Table 2.1.: [Raabe: Molecular Simulation Studies on Thermophysical Properties; Springer 2017] Comparison of the classical ensemble types: the control variables by which they are characterized and the properties that fluctuate.

ensemble		control variables	important fluctuating quantities	corresponding thermodynamic system
microcanonical (NVE)		number of molecules N volume V internal energy E	pressure energy of each molecule ϵ_i	isolated
canonical (NVT)		number of molecules N volume V temperature T	pressure energy E	closed isochoric thermostated
grand canonical (μVT)		chemical potential μ volume V temperature T	pressure no. of molecules N energy E	open with connection to substance-reservoir thermostated
isothermal-isobaric (NPT)		number of molecules N pressure P temperature T	volume V energy E	closed isobaric isothermal

simulation and therefore, in principle, is a good choice for the determination of vapor liquid equilibria (VLE). The details and techniques to achieve good results for VLEs will be discussed in section 2.2.

There is one more important ensemble to mention, especially for the determination of phase equilibria, in particular VLE. The Gibbs ensemble proposed by Panagiotopoulos [69] was developed to directly simulate vapor-liquid equilibria. The total volume is split in two sub-volumes without physical contact to each other, however, molecules can be removed from one sub-volume and inserted in the other and the size of each sub-volume can change but only in such a way that the total volume stays constant. This arrangement leads to equilibrium properties where one sub-volume represents the vapor phase and the other sub-volume the liquid phase. The pressure and the densities can be determined through ensemble averages. To get a complete phase diagram many independent simulations at different conditions have to be done and evaluated.

Despite the classical usages of the ensembles above, a lot of different approaches and methods have been developed over the last decades and years to describe and predict phase equilibria with molecular simulations and post-processing. For a comprehensive overview I recommend [8–12, 69–71].

2.2. Grand Canonical Monte Carlo simulations to determine phase equilibria

In classical thermodynamics, three conditions need to be met for a thermodynamic system to be in equilibrium state. The three conditions are equality of the temperature $T^I = T^{II}$, mechanical equilibrium $p^I = p^{II}$ and chemical equilibrium $\mu_i^I = \mu_i^{II}$ for every substance i in a mixture. If more than two phases are in equilibrium, the equality conditions have to be fulfilled for all phases. I will focus on the Grand Canonical ensemble and its advantages for the purpose of this work.

A big advantage of simulations in the Grand Canonical ensemble is that the temperature T and chemical potentials μ_i are control variables and are thus user-defined. If suitable values for these quantities are chosen, it is possible to determine two phases (in our case liquid and vapor phase) in a single simulation. Because the box volume V of the simulation is the third fixed control parameter, the number of molecules has to fluctuate during a simulation. Because the exact chemical potentials $\mu_i(T, V)$ that lead to coexisting phases for a given force field and for defined temperature T and volume V are a priori unknown, estimates for chemical potentials for conducting the simulations are used. Estimates are sufficient because histogram reweighting techniques can be applied for relating the simulated histograms to histograms at other chemical potentials. The chemical potentials that ensure coexisting phases can then be determined as a post-processing step[1, 6, 72–75]. Histogram reweighting requires sampling histograms, i.e. probability distribution functions of type $p(N, E)$ for defined $\bar{\mu}, V, T$. Details to these technique are described in section 2.2.2.

Figure 2.1 schematically shows a projection of this probability distribution $\Pi(N) = \sum_E p(N, E)$, for constant μ, V, T somewhere between the triple point and the critical point of a pure substance. The scheme shows two peaks with a high probability of finding N particles. The peak at low number of molecules represents the vapor phase and the peak at higher N corresponds to the liquid phase. The “area” under each peak corresponds to the pressure of this phase. If the chemical potentials are not exactly equal to the equilibrium values $\bar{\mu}^{\text{coex.}}$, then the areas are not the same and the system is not in its equilibrium state.

The space between the two peaks belongs to a region that would for a macroscopic system be meta-stable or unstable and has a very low probability. Low probabilities are challenging for Grand Canonical Monte Carlo (GCMC) simulations, as the transfer between the vapor and the liquid phase has to happen but is very unlikely. This is a challenge, because for statistical reasons the number of samples in the entire range of molecule number N should be approximately equal. If the region between the two phases is too unlikely, the transition will never happen without introducing a bias.

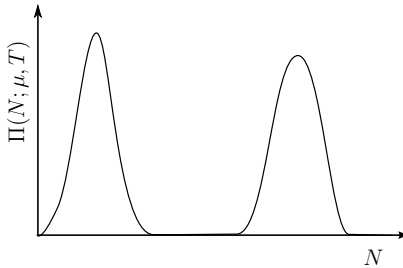


Figure 2.1.: Schematical depiction of the unbiased probability distribution $\Pi(N; \mu, T)$.

Luckily, this issue can be solved by applying one of several sampling techniques. A first simple measure for improving the sampling is to divide the N -space into several windows of size $\Delta_k N$, where each window k can be simulated independently of each other. However, splitting the N -space into several windows is often not sufficient to overcome the regions of low probability between the two phases. Therefore, techniques like Transition Matrix Monte Carlo and Multiple Histograms are necessary. These techniques ensure that the complete phase space is sampled evenly and they allow us to split a simulation into several well-sampled parts. Those parts can furthermore be conducted for different conditions μ_k, T_k and can run in parallel [1, 6]. Details are given in the following section 2.2.1

2.2.1. Transition matrix Monte Carlo and multiple histograms

To determine the phase equilibria between two phases the probability distribution is of key interest. As the name implies, this distribution shows the probability of finding the system with a certain density (or a certain number of particles). Since the probability of finding a system between two phases (macroscopically speaking, in the meta-stable region) is extremely small, one has to apply special methods to overcome the gap between these phases. All of the following techniques target at flat histograms where all states should be sampled uniformly by introducing bias functions.

In a histogram different states are collected by defining bins of a certain size and by counting how often the system visits every bin. For Monte Carlo simulations of pure substances histograms are two-dimensional matrices $H(N, E)$, where the number of molecules N takes discrete integer values and the energy is split into different bins of size ΔE . If a binary mixture is considered, the histogram is three-dimensional $H(N_1, N_2, E)$. To achieve flat histograms a biasing function has to be introduced and the flat histogram can be achieved by step-wise updating the biasing function until a convergence criterion is reached. To calculate the biasing function either all visited states can be counted or - which increases the accuracy enormously - the transition probability of each step is taken

2. Theory and Methods

into account. The latter is called the transition matrix method. The main idea of the transition matrix technique is described by [Fitzgerald et al. \[76, 77\]](#). In their publications they describe a simple bookkeeping step which enables them to obtain statistical estimators of canonical probabilities (referred to as macrostate probabilities in the publication by Fitzgerald et al.). The use of Boltzmann transition probabilities in combination with a simulation with non-Boltzmann sampling methods increases the obtained information significantly. They demonstrated their method with the calculation of the interfacial tension and magnet susceptibility of the 2D Ising system. [Errington and Coworkers \[78, 79\]](#) transferred this approach to Monte Carlo simulations in grand canonical and isothermal-isobaric ensembles to directly calculate coexistence properties in a temperature range from the triple point to the critical point. Following this approach the probability $\pi(N, E)$ with N of being in a microstate molecules and the total energy E in grand canonical ensemble is

$$\pi(N, E) = \frac{1}{\Xi} \frac{V^N}{\Lambda^{3N} N!} \exp[-\beta E + \beta \mu N]. \quad (2.1)$$

The probabilities of microstates with $N + 1$ or $N - 1$ particles correspondingly is

$$\pi(N + 1, E) = \frac{1}{\Xi} \frac{V^{N+1}}{\Lambda^{3(N+1)} (N + 1)!} \exp[-\beta E + \beta \mu (N + 1)] \quad (2.2)$$

$$\pi(N - 1, E) = \frac{1}{\Xi} \frac{V^{N-1}}{\Lambda^{3(N-1)} (N - 1)!} \exp[-\beta E + \beta \mu (N - 1)] \quad (2.3)$$

and the standard acceptance criterion says

$$p_{s \rightarrow t}^{\text{acc}} = \min \left[1, \frac{\pi(t)}{\pi(s)} \right] \quad (2.4)$$

where s is the current microstate and t indicates a microstate that is to be reached by a random move. Inserting the probabilities from above with $s \doteq N$ and $t \doteq N + 1$ leads to the acceptance criteria for insertion and deletion

$$p_{N \rightarrow N+1}^{\text{acc}} = \min \left[1, \frac{V}{\Lambda^3 (N + 1)} \exp(-\beta \Delta E + \beta \mu) \right] \quad (2.5)$$

$$= \min \left[1, \frac{V}{(N + 1)} \exp(-\beta \Delta E + \beta \mu^{ex}) \right] \quad (2.6)$$

$$p_{N \rightarrow N-1}^{\text{acc}} = \min \left[1, \frac{N}{V} \exp(-\beta \Delta E - \beta \mu^{ex}) \right], \quad (2.7)$$

with $\beta \mu^{ex} = \beta \mu - \ln \Lambda^3$

Figure 2.1 shows the unbiased probability distribution for a system at given chemical potential μ and temperature T . Since the probability of being in a state between the

2. Theory and Methods

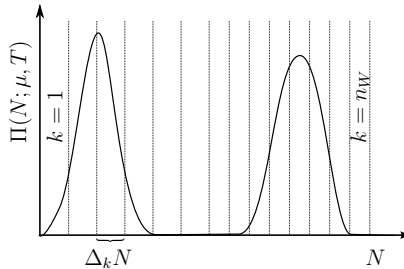


Figure 2.2.: Schematical depiction of the probability distribution $\Pi(N; \mu, T)$ and the division into several windows k with window size $\Delta_k N$.

vapor and liquid peak is extremely small, normal acceptance rules would keep the simulation either in the vapor state, or in the liquid state. For the best statistical results it would be ideal, if all states had the same probability and therefore the distribution was a horizontal line. To achieve this bias function is introduced which is exponential of the reciprocal function of the estimated probability distribution during a GCMC simulation. Using this a bias function $\omega(N)$, the resulting biased acceptance probabilities are

$$a_{N \rightarrow N+1} = \min \left[1, \frac{V}{N+1} \exp(-\beta \Delta E + \beta \mu^{\text{ex}} + \Delta \omega) \right] \quad (2.8)$$

$$a_{N \rightarrow N-1} = \min \left[1, \frac{N}{V} \exp(-\beta \Delta E - \beta \mu^{\text{ex}} + \Delta \omega) \right] \quad (2.9)$$

where $\Delta \omega = \omega_t - \omega_s$ (and analogously for ΔE and $t \hat{=} N + 1$ and $s \hat{=} N$). The transition matrix approach delivers both, a way to estimate the bias function $w(N)$ such that the probability of visiting states with N molecules is flat, and it allows for recovering an estimate of the non-biased probability distribution. In transition matrix a bookkeeping step is introduced to collect all unbiased acceptance probabilities in a collection matrix C . This collection matrix is – for a pure substance – a tridiagonal matrix. For each particle number N – besides 0 and N_{max} – the matrix C contains an entry for the accumulated probability of moving to a state with $N - 1$, $N + 1$ particles or staying in a state with N particles. The iteration instruction to update the collection matrix C for an insertion move is

$$C_{N \rightarrow N+1} \doteq C_{N \rightarrow N+1} + a_{N \rightarrow N+1} \quad (2.10)$$

$$C_{N \rightarrow N} \doteq C_{N \rightarrow N} + (1 - a_{N \rightarrow N+1}) \quad , \quad (2.11)$$

where $a_{N \rightarrow N+1}$ is the unbiased acceptance probability.

2. Theory and Methods

For displacement, regrowth and rotation moves the instruction is

$$C_{N \rightarrow N} \doteq C_{N \rightarrow N} + 1. \quad (2.12)$$

With the entries of the collection matrix C the unbiased transition probability between macrostates can be estimated

$$P(N \rightarrow N + \delta) = \frac{C(N \rightarrow N + \delta)}{C(N \rightarrow N - 1) + C(N \rightarrow N) + C(N \rightarrow N + 1)} \quad (2.13)$$

with $\delta = \{-1, 0, 1\}$ and the unbiased probability of being in a state can be determined by

$$\Pi(N + 1) = \Pi(N) \cdot \frac{P(N \rightarrow N + 1)}{P(N + 1 \rightarrow N)} \quad (2.14)$$

Setting $\Pi(N = N_k^{\min}) = 1$ leads to an unnormalized probability distribution, which at this point is sufficient. To obtain an almost flat histogram the bias function $\omega(N)$

$$\omega(N) = -\ln(\Pi(N)) \quad (2.15)$$

is introduced.

While the simulation is carried out with the biased acceptance criteria, the entries for the collection matrix C are updated with the unbiased values. To finally determine the

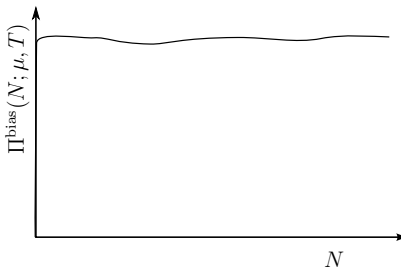


Figure 2.3.: Schematical depiction of the biased probability distribution $\Pi^{\text{bias}}(N; \mu, T)$.

phase coexistence points not only the probability distribution as a function of N is needed, but also a distribution as function of N and the energy E . This is done with a histogram $H(N, E)$ where the energy landscape is divided into bins and it is counted how often the system energy lies within each bin. Since a bias was added to the acceptance criteria, the resulting histogram is of type $H^{\text{bias}}(N, E)$.

Additionally to the transition matrix method, the phase space of N can be divided into several windows $k = 1 \dots n_W$ of size $\Delta_k N$, as already shown in Figure 2.2. Figure 2.4

2. Theory and Methods

again shows the divided phase space, however, this time several probability distributions at different conditions are added. The black dotted line is for a low temperature where a liquid and a vapor phase exists. The blue solid line is at a higher temperature and the red dash-dotted line is at the highest temperature, near the critical point. Obviously, the peaks at different conditions appear at different particle numbers and the gap between the two phases gets smaller towards the critical region. Since the quality of the result depends on the accurate sampling of these probability functions, it is useful to sample the different windows at different conditions. All shown probability functions (here black, blue and red) can afterwards be calculated with the histogram reweighting method, although the different windows have been simulated with different conditions. To get an estimation

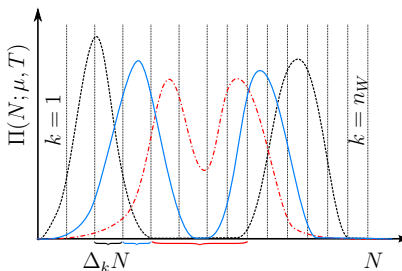


Figure 2.4.: Schematical depiction of the probability distribution $\Pi(N; \mu, T)$ and the division into several windows k with window size $\Delta_k N$ at different conditions μ, T , where the black dotted curve is at the lowest temperature and the red dash-dotted line at the highest temperature near the critical point.

of plausible conditions at which the GCMC simulations for domains of different particle numbers in the system should be carried out, the PC-SAFT equation of state is used. Figure 2.5 schematically shows these different conditions along a phase envelope. During the simulation, histograms of kind $H_i^{\text{bias}}(N, E)$ will be collected for all different conditions i . In detail, this works in the order of the following steps. The width of the energy bins ΔE is given as an input parameter to the simulation. It is also a parameter which at the end of this work was seen to have an influence on the results for some substances at lower temperatures (see Appendix D). Therefore, the values of ΔE will be reported in every result. A discussion of this behavior is given in Appendix D.

As a post-processing step, after the production cycle is finished, the unbiased probability distribution is calculated from (2.11) and using [1]

$$H(N, E) = \frac{H^{\text{bias}}(N, E)}{\sum_E H^{\text{bias}}(N, E)} \cdot \Pi(N) \quad (2.16)$$

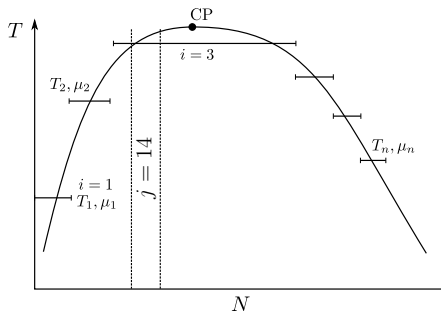


Figure 2.5.: Different simulation conditions along the phase envelope for simulations in grand canonical ensemble [1].

Afterwards, the neighboring windows $H_i(N, E)$ and $H_{i+1}(N, E)$ are combined at the overlapping particle number $N_{i+1}^{\min} = N_i^{\max}$. This step is trivial for neighboring windows with the same conditions μ, T and a bit more sophisticated for different conditions where histogram reweighting has to be applied. Section 2.2.2 describes this step for neighboring windows at different conditions. The sampling of different windows at different conditions which can be combined afterwards is referred to as multiple histograms.

2.2.2. Histogram reweighting for pure substances

With Histogram reweighting it is possible to obtain information from an energy probability function $H_i(N, E) = H(N, E; \beta_i, \mu_i)$ at the condition β_i, μ_i for a different condition β, μ , with $H(N, E) = H(N, E; \beta, \mu)$. The original method was developed by [Ferrenberg and Swendsen](#) [72, 73] for a 2D Ising model. [Panagiotopoulos](#) [13] gave an excellent review of this technique as used in GCMC simulations. A short overview of this technique and its simplifications is given for the cases studied in this work. The probability distribution $H(N, E)$ at given β, μ follows the relationship

$$H(N, E; \beta, \mu, V) = \frac{\Omega(N, V, E) \cdot \exp(-\beta E + \beta \mu N)}{\Xi(\mu, V, \beta)} \quad (2.17)$$

where $\Omega(N, V, E)$ is the microcanonical partition function and $\Xi(\mu, V, \beta)$ the grand partition function. The left hand side of equation (2.17) is an output of simulations in grand canonical ensemble. The microcanonical partition function $\Omega(N, V, E)$ contains the integration over all configurations and is independent of β and μ . Had we conducted a simulation at a different value of β_i and μ_i , where (N, E) is an accessible and visited state, we have the same value of $\Omega(N, V, E)$. Formulating equation (2.17) for a probability

2. Theory and Methods

distribution at different conditions β_i, μ_i , setting $\Omega(N, V, E)$ equal (and thus setting the entropies equal) and rearranging the resulting equation leads to

$$H(N, E; \beta, \mu) = H_i(N, E; \beta_i, \mu_i) \cdot C_i \cdot \exp[-(\beta - \beta_i)E + N(\beta\mu - \beta_i\mu_i)] \quad (2.18)$$

with the constant C_i being the fraction of the grand partition functions

$$C_i = \frac{\Xi(\mu_i, V, \beta_i)}{\Xi(\mu, V, \beta)} \quad (2.19)$$

which is not known a priori. In the original articles on histogram reweighting, these constants C_i are determined iteratively and using weight-functions to account for statistical uncertainties, the different overlapping areas[72–74]. The transition matrix scheme used in this work, however, ensures sufficient sampling of two neighboring simulations, one at β, μ and one at β_i, μ_i for the same (overlap) molecule number N_{ov} . We require the probabilities are identical for averaged $\bar{\beta}$ and average chemical potential $\bar{\beta}\bar{\mu}$ and the overlap occurs at one particle number N_{ov} . The Histograms can be determined by

$$H_{i+1} \doteq H_{i+1} \frac{\sum_E H_i \exp[-(\bar{\beta} - \beta_i)E + N_{ov}(\bar{\beta}\bar{\mu} - \beta_i\mu_i)]}{\sum_E H_{i+1} \exp[-(\bar{\beta} - \beta_{i+1})E + N_{ov}(\bar{\beta}\bar{\mu} - \beta_{i+1}\mu_{i+1})]} \quad (2.20)$$

with $H_i = H_i(N_{ov}, E; \beta_i, \mu_i)$ and $H_{i+1} = H_{i+1}(N_{ov}, E; \beta_{i+1}, \mu_{i+1})$. Applying this procedure for all conditions i leads to a non-normalized energy distribution function $H(N, E; \beta, \mu)$.

The probability distribution $P(N; \beta, \mu)$ is then calculated by summing up over all energy bins of the energy distribution function and by normalizing it according to

$$P(N; \beta, \mu) = \frac{\sum_E H(N, E; \beta_j, \mu_j) \exp(-(\beta - \beta_j)E + N(\beta\mu - \beta_j\mu_j))}{\sum_N \sum_E H(N, E; \beta_j, \mu_j) \exp(-(\beta - \beta_j)E + N(\beta\mu - \beta_j\mu_j))}. \quad (2.21)$$

with j being a generic index. If, for example the 14th simulation (see fig. 2.5) covers molecule numbers from $N = 130$ to 140, then index $j = 14$ for all molecule numbers $130 \leq N < 140$ in equation (2.20).

The determination of the phase equilibrium also works with histogram reweighting. As mentioned before, the requirements are equality of temperature, chemical potential and pressure. The first two conditions are trivially fulfilled through the choice of the GCMC ensemble where histogram reweighting leads to equations for the probability distribution function as a function of temperature and chemical potential. The equality of pressure is given, if the “area under each peak” (fig. 2.1), or more precisely the sum over all probabilities for all energies and molecule numbers are equal $I^V = I^L$. The two sums

2. Theory and Methods

write as

$$I^V = \sum_{N=0}^{N^{\text{sep}}} P(N; \beta, \mu^{\text{coex}}(\beta)) \quad (2.22)$$

$$I^L = \sum_{N=N^{\text{sep}}}^{N^{\text{max}}} P(N; \beta, \mu^{\text{coex}}(\beta)) \quad (2.23)$$

and N^{sep} is the number of particles where $P(N; \beta, \mu^{\text{coex}}(\beta))$ has its minimum between the two phases. The histogram reweighting is done for different temperatures T along the phase envelope. For each $T = 1/(k\beta)$ the chemical potential $\mu^{\text{coex}}(\beta)$ is determined by a Newton algorithm to find the zero of the function $f = \ln(I^v/I^l)$. Of course, if $\ln(I^v/I^l) = 0$ it follows $I^v = I^l$. As soon as the iteration sufficiently converged (usually in three to five steps) and the phase coexisting point for the given temperature T is therewith determined, the thermodynamic values can be calculated from the probability distribution at this equilibrium point $P^{\text{coex}}(N, T) = P(N, T; \mu^{\text{coex}}(T))$. The ensemble averages then are

$$\langle A \rangle^v = \sum_{N=0}^{N^{\text{sep}}} 2P^{\text{coex}}(N, T)A(N) \quad (2.24)$$

$$\langle A \rangle^l = \sum_{N=N^{\text{sep}}}^{N^{\text{max}}} 2P^{\text{coex}}(N, T)A(N). \quad (2.25)$$

The vapor pressure of the coexistence point is calculated with $\ln \Xi = \beta pV$ in the ideal gas limit at low densities [6, 13] where the system is assumed to follow the ideal gas equation of state $pV = Nk_B T$

$$p^{\text{sat}}(T) = -\frac{k_B T}{V} \ln(2P^{\text{coex}}(N=0; T)). \quad (2.26)$$

All equations and the iteration of μ towards $\mu^{\text{coex}}(T, V)$ are performed at various temperatures in order to determine a full vapor-liquid phase diagram of pure substances.

2.2.3. Histogram Reweighting for binary mixtures

For binary mixtures, the dimension of the collected histogram expands to three dimensions. As computer power and storage increased rapidly over the last twenty years - and are still increasing - this, in contrast to the review given by Panagiotopoulos[13], is no longer a problem. Several different studies in the past used simulations in the grand canonical ensemble combined with histogram reweighting methods. Studies from the late 1990s where the critical points of a pure Lennard-Jones fluid and a Lennard-Jones mixture were investigated[80], as well as a study for mixtures of polar and non-polar components[81], stored a list of N_1 , N_2 and E -values in every 250th step and extracted

the probability distribution afterwards, which reduced the amount of necessary storage significantly. Later, this procedure was applied to more sophisticated force fields with a special focus on the behavior of $CO_2 - H_2O$ -mixtures[82] for different kinds of force fields for CO_2 and water and the different combinations of these force fields in mixtures. In all of these studies they needed to do the first simulation near the critical point of one component, to then reweight the results to lower temperatures where additional simulations were conducted. In this thesis, simulations near the critical point were not necessary as a starting point, because a physically based equation of state was used to get a sufficient initial guess of the input properties of the simulations and all desired simulations could run in parallel and be combined in the end. This procedure is described in chapter 2.3.

Combining histograms of various windows to a single Molecule Number Probability Distribution

For binary mixtures the entire N_1 - N_2 -grid is divided into various windows of size $\Delta N_1 \times \Delta N_2$. A probability distribution can be determined for each of these windows. To get an overall probability distribution it is necessary to combine the probability distributions of all windows at the overlapping windows-boundaries. The probability $\Pi(\bar{N} = \bar{0}; \bar{\mu}, v, T)$ is initially set to unity and the normalization is done after combining all probability distributions. All probability distributions will then be combined according to

$$\Pi(\mathbf{N}; \boldsymbol{\mu}, V, T) = \Pi^W(\mathbf{N}; \boldsymbol{\mu}, V, T) \cdot \exp(\Gamma), \quad (2.27)$$

where $\Pi(\mathbf{N}; \boldsymbol{\mu}, V, T)$ is the probability distribution containing all already combined windows, the superindex W corresponds to the currently considered window and $\exp(\Gamma)$ is the scaling factor, that is needed to combine the current window with all other windows. Three different cases have to be taken into account. If the window contains a pure component 2 ($N_2 = 0$) it has an overlap only on the left hand side to another window, as illustrated in Fig. 2.6. Analogously, if the window contains a pure component 1 ($N_1 = 0$) then it has an overlap only at the “bottom”, meaning towards lower N_2 -values. For all other cases the window has an overlap at two boundaries (see figure 2.6). There are different options to calculate the scaling factor. The most intuitive way would be to start along the axis N_1 and N_2 and to calculate Γ according to the average offset along the corresponding overlap:

2. Theory and Methods

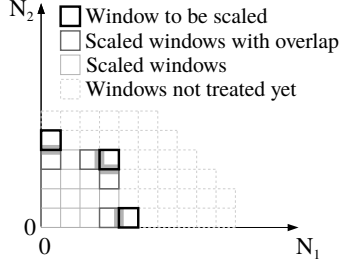


Figure 2.6.: Possible overlapping cases to determine the scaling factor Γ .

$$\Gamma_1 = \frac{\sum_{N_b=N_2^{W,\min}}^{N_2^{W,\max}} \ln \Pi \left(N_1^{W,\min}, N_b; \boldsymbol{\mu}, V, T \right) - \ln \Pi^W \left(N_1^{\min}, N_b; \boldsymbol{\mu}, V, T \right)}{N_2^{W,\max} - N_2^{W,\min}} \quad (2.28)$$

$$\Gamma_2 = \frac{\sum_{N_a=N_1^{W,\min}}^{N_1^{W,\max}} \ln \Pi \left(N_a, N_2^{W,\min}; \boldsymbol{\mu}, V, T \right) - \ln \Pi^W \left(N_a, N_2^{\min}; \boldsymbol{\mu}, V, T \right)}{N_1^{W,\max} - N_1^{W,\min}}. \quad (2.29)$$

However, for the case where the windows are at the left bound ($N_1^{\min} = 0$) or at the bottom bound ($N_2^{\min} = 0$), the phase diagrams will get more accurate if the connecting factor Γ_r between two windows is calculated as

$$\Gamma_p = \ln \Pi \left(N_1^{W,\min}, N_2^{W,\min}; \boldsymbol{\mu}, V, T \right) - \ln \Pi^W \left(N_1^{\min}, N_2^{W,\min}; \boldsymbol{\mu}, V, T \right). \quad (2.30)$$

This means, only the overlap at the lower left corner of the currently considered window is used to calculate the scaling factor. Along both axis this is equivalent to patching the windows in such a way that the resulting overall probability distribution is smooth for the pure substances. This is not necessarily the case if the mean value along the overlapping side is considered. Finally, this leads to a more accurate description of the vapor pressure for both pure substances and in consequence, the borders of the resulting phase diagram at constant temperature are more accurate. Figures in the chapter “Results” show the influence of the different treatments of patching the histograms. For all windows that do not contain a pure substance the scaling factor is calculated along both overlapping lines. Since an arithmetic mean of Γ_1 and Γ_2 would take the lower left corner of the windows

2. Theory and Methods

into account twice, the scaling factor is calculated as $\Gamma = A \cdot \Gamma_A + B \cdot \Gamma_B$, with

$$A = \frac{N_2^{W,\max} - N_2^{W,\min}}{N_2^{W,\max} - N_2^{W,\min} + N_1^{W,\max} - N_1^{W,\min} - 1} \quad (2.31)$$

$$B = \frac{N_1^{W,\max} - N_1^{W,\min} - 1}{N_2^{W,\max} - N_2^{W,\min} + N_1^{W,\max} - N_1^{W,\min} - 1} \quad (2.32)$$

and

$$\Gamma_A = \frac{\sum_{N_b=N_2^{W,\min}}^{N_2^{W,\max}} \ln \Pi(N_1^{W,\min}, N_b; \boldsymbol{\mu}, V, T) - \ln \Pi^W(N_1^{\min}, N_b; \boldsymbol{\mu}, V, T)}{N_2^{W,\max} - N_2^{W,\min}} \quad (2.33)$$

$$\Gamma_B = \frac{\sum_{N_a=N_1^{W,\min}+1}^{N_1^{W,\max}} \ln \Pi(N_a, N_2^{W,\min}; \boldsymbol{\mu}, V, T) - \ln \Pi^W(N_a, N_2^{\min}; \boldsymbol{\mu}, V, T)}{N_1^{W,\max} - N_1^{W,\min}}. \quad (2.34)$$

The difference of Γ_A and Γ_B to Γ_1 and Γ_2 is subtle. It manifests only in the lower bound of the sum in Γ_B compared to Γ_2 , while Γ_A equals Γ_1 .

2.2.4. Configurational bias Monte Carlo (CBMC)

Considering the set of Monte Carlo moves in a grand canonical ensemble: The different moves which can be performed are translation, rotation, insertion, deletion and rebuilding a molecule. Especially for higher densities, the probability of accepting an insertion or deletion move (or a rebuilding move) decreases rapidly, because for many trial configurations one is faced with high energies and thus low probabilities of accepting trial moves. An efficient way to improve the acceptance probability is called configurational bias (CB) and was proposed by [Siepmann and Frenkel](#) [83] and extended to branched molecules by [Martin and Siepmann](#)[51] as well as [T.J.H. Vlugt and Krishna](#)[84] and is mainly based on works of [Rosenbluth and Rosenbluth](#) [85]. With configurational bias, a molecule is successively grown (or regrown), interaction site by interaction site, with different trial positions for every interaction site. During this growth procedure the probabilities are determined and summed up and finally a configuration is selected as a trial configuration according to its Rosenbluth weight

$$W_n = \prod_{f=1}^M w_f \quad (2.35)$$

2. Theory and Methods

Table 2.2.: Number of CB steps for different dimensionless densities $\rho^* = \rho/\rho_{\max}$

ρ^*	CB steps per interaction site
$< 1/2.7$	1
$\geq 1/2.7 \wedge < 1/1.9$	2
$\geq 1/1.9 \wedge < 1/1.6$	3
$\geq 1/1.6 \wedge < 1/1.3$	4
$\geq 1/1.3 \wedge < 1/1.16$	6
$\geq 1/1.16$	8

where w_f is – analogously to a partition function – the sum over all k different probabilities of trial positions for the actual CB step (index f), as

$$w_f = \sum_{e=1}^k \exp[-\beta u_e^{\text{inter}}] \quad (2.36)$$

where u_e^{inter} covers all intermolecular interactions and the probability for selecting all sites at their trial position e is

$$p_e = \frac{\exp[-\beta u_e^{\text{inter}}]}{w_f}. \quad (2.37)$$

The acceptance probability $p_{o \rightarrow n}^{\text{acc}}$ for finally accepting the trial move (where all interaction sites $f = m, \dots, M$ have been placed) then writes

$$p_{o \rightarrow n}^{\text{acc}} = \min\left(1, \frac{W_n}{W_o}\right) \quad (2.38)$$

with W_n and W_o being the Rosenbluth weights of the trial (new) and old configuration of a molecule, respectively. This procedure can be applied to the regrowth of a molecule as well as for insertion and deletion moves. In this work the same number of CB steps is applied for all sites to be (re)located and this number is dependent on the density of the considered system. For lower densities near the ideal gas state the acceptance probability for a move is often good enough to omit additional configurational bias. For high densities, e.g. the liquid state at temperatures near the triple point, the number of CB steps is set to 8. Table 2.2 lists the number of CB steps depending on the dimensionless density as used in our simulations.

2.2.5. A fragment library for branched and cyclic molecules

For sites with a “linear” connection scheme, like in n -alkanes or primary alcohols the intramolecular trial configurations entering the CBMC algorithm can successively be determined. For branched or cyclic structures that is different. If one imagines the branching

2. Theory and Methods

point of a molecule – that means one interaction site is bonded to at least 3 neighboring interaction sites – the challenge is to get the correct angle distribution. CBMC for a “linear” connection scheme is executed interaction site by interaction site, which in this case would lead to a biased angle distribution [51, 86].

An elegant method to generate correct angle distributions is to sample each branch on its own (consisting of 3 or more neighboring interaction sites around a central “interaction site”, i.e. a branch does not contain a torsional angle potential. Afterwards, the branch will be connected to the existing molecule and only the torsional energy determines if the move will be accepted or not. This approach is based on works of Shah and Maginn [87] and Errington and Panagiotopoulos [88], who took the idea of the branch-point sampling (BPS) by Macedonia and Maginn [89] and extended it to be more efficient. BPS is based on the idea that there are hard and soft degrees of freedom which can be sampled independently of each other. The basically same idea, namely that the energy of hard degrees of freedom – like bending angles of a branched molecule – is sensitive to small changes in interaction site positions, while the energy of soft degrees of freedom is weaker dependent on interaction site positions was proposed in the same year by Martin and Siepmann [51]. In their work Shah and Maginn generate the repository of configurations *a priori* to the actual Monte Carlo simulation according to the given temperature of the following simulation.

In this work a slightly different method is chosen and the repository (or library) is generated during runtime, which means for the first, say 100,000, insertions or regrowth steps a mini Monte Carlo simulation is done to generate a configuration of a branch or molecule. This configuration is then stored in the library which is a simple array of relative positions. After the first 100,000 steps the array is completely filled and all following regrowth or insertion steps are based on the configurations stored in the library. For branched molecules, in a first step, a branch is taken randomly from the library and added to the existing part of the molecule by putting two interaction sites along a bonding axis so that they overlap with two interaction sites of the already grown part (see Figure 2.7). In the second step different trial positions are generated by rotating the new branch around the bonding axis. This second step can be handled with configurational biasing methods, as discussed in section 2.2.4. For cyclic molecules the challenge is to “lead” the configurational biasing steps in such a way, that the ring closes with the insertion of the last interaction site and with the correct bond lengths. The Self-Adapting Fixed-Endpoint (SAFE) CBMC method by Wick et al. [90] solves exactly this problem. Nonetheless, this sophisticated sampling scheme is not easy to implement and for small cyclic molecules the use of a fragment library leads to the same results with less effort. For cyclopentane or cycloheptane e.g. a branch consists simply of the whole molecule and the different intramolecular configurations can be generated with crank-shaft moves [91, 92].

2. Theory and Methods

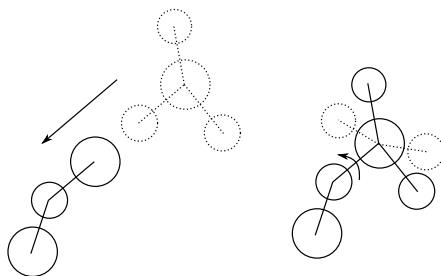


Figure 2.7.: Schematic illustration of a growing step with the fragment library. On the left-hand side of the figure a branch from the library is placed over the existing bond axis of the already grown part of a molecule. On the right-hand side different trial configurations are generated by rotating the branch around the axis.

Figure 2.8 shows the generation of configurations for cyclohexane with crank-shaft moves.

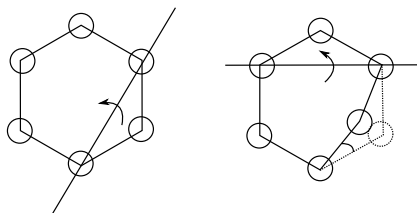


Figure 2.8.: Schematic depiction of crank-shaft moves for a cyclic molecule with six interaction sites.

Randomly three adjacent interaction sites are chosen and the middle interaction site is rotated with a random angle around the axis through its neighboring interaction sites. The new configuration is accepted according to standard Metropolis acceptance criterion and the resulting distribution is a Boltzmann distribution. The standard configurational biasing methods can be applied for the first three interaction sites of a cyclic structure where the position of the third interaction site has to fulfill the constraint (the angle between the first three interaction sites) given by the randomly chosen library fragment. For small molecules like cyclopentane or cyclohexane this procedure leads to good acceptance probabilities and hence to sufficient sampling.

2.3. The PC-SAFT equation of state and its application to force field optimization

The following sections describe the Perturbed Chain Statistical Associating Fluid Theory (PC-SAFT) and its use for generating different state conditions along the phase envelope as an input for GCMC simulations (section 2.3.2) and its use for accelerating the optimization of TAMie force field parameters.

2.3.1. The perturbed chain statistical associating fluid theory (PC-SAFT) equation of state

The Perturbed Chain Statistical Associating Fluid Theory developed by Gross and Sadowski [59–61] is based on Wertheim’s perturbation theory [93–96] and constructed in terms of the Helmholtz free energy. Like other SAFT equations it is based on works of Chapman et al. [97, 98]. Over the years several modifications like VR-SAFT [99], SAFT- γ [57, 100], LJ-SAFT [101] and PC-SAFT [59–61] were developed.

Starting from hard chains as a reference fluid a perturbation theory of second order as proposed by Barker and Henderson [102, 103] was applied to get the dispersive part for chain molecules. All contributions, like dispersive interactions or associating terms, are additive with respect to the compressibility factor Z or the (dimensionless) Helmholtz free energy $\hat{a} = A/NkT$.

$$\hat{a} = \frac{A}{NkT} = \frac{A^{\text{hc}}}{NkT} + \frac{A^{\text{disp}}}{NkT} + \frac{A^{\text{assoc}}}{NkT} + \frac{A^{\text{polar}}}{NkT} \dots \quad (2.39)$$

The PC-SAFT model has the following contributions: the hard chain contribution plus the dispersive interactions [59], attractive contributions due to hydrogen bonding [61] and electrostatic interactions [62, 104, 105].

One of the main difference of PC-SAFT to other versions of SAFT is the choice of hard chains as reference fluid. The hard-chain reference was developed by Chapman et al. [97, 98] based on Wertheim’s [93–96] thermodynamic perturbation theory of first order. The underlying molecular model for the chain fluid is that of hard spheres which are tangentially bound. The PC-SAFT equation of state has different variants of parameterization. One variant is the parameterization for pure components, where each substance gets its unique set of parameters. The other variants are group contribution methods, where a chemical group, e.g. a CH_3 group in alkanes, has its set of parameters and all molecules of the alkanes series share the same group parameters but differ in the number of groups. The last differentiation is based on how the chain is build. In the homo-segmented group contribution method the final molecule has m segments and each

segment has the same parameters σ and ε , whereas in the hetero-segmented group contribution method the single segments can differ in size and interaction parameters. For this work, the homo-segmented group contribution method and PC-SAFT for pure substances was used.

2.3.2. Generating conditions for GCMC simulations with PC-SAFT

Like all thermodynamic equations of state, the PC-SAFT equation of state (eos) links the state variables (T, p, V, N & residual properties) of a thermodynamic system and can be used to determine phase equilibria. The PC-SAFT eos is formulated in terms of the Helmholtz free energy and it is easy to calculate phase coexistence points for different specified temperatures with the according residual chemical potentials μ_k^{res}

$$\mu_k^{\text{res}}(T, v) = k_B T \left(\frac{\partial(\rho \hat{a}^{\text{res}})}{\rho_k} \right)_{T, v, \rho_j \neq k}. \quad (2.40)$$

The PC-SAFT eos in this work is used to support molecular simulations by defining suitable (T, μ) -conditions approximately tracing the phase envelope of a substance, as shown in Fig. 2.5. If we consider acetone, for example, the phase envelope calculated by PC-SAFT approximates the experimental behavior of acetone rather well. If a force field used for acetone is also roughly reproducing the phase behavior the (T, μ) -conditions of PC-SAFT will roughly approximate the relevant (T, μ) -conditions for the force field. Because histogram reweighting is used, (rough) estimated (T, μ) -conditions are sufficient. The eos requires pure component parameters or group-contribution parameters of the considered molecule as input.

2.3.3. Optimization of Mie parameters with PC-SAFT

Classical optimization of force field parameters for molecular simulations need two runs with different sets of parameters to receive gradients from which the next set of parameters follow. The procedure of calculating gradients from different simulations is intuitive but slow, because each molecular simulation needs quite a lot of time and even for the first guess two simulations are necessary. Another possibility to optimize force field parameters is the method to reweight the parameters from one simulation to a nearby parameter space. The Multistate Bennett Acceptance Ratio (MBAR) [106] developed by [Shirts and Chodera](#) is such a method which allows the reweighting to different conditions (compareable to the histogram reweighting method) or to different parameters. The method to reweight the parameters with MBAR is constrained to a small variations in parameter space near the original simulation.

2. Theory and Methods

van Westen et al. showed in 2011 that Mie parameters of molecular force field can be approximated with modified PC-SAFT parameters. In 2015 this method was improved by Hemmen et al.. This enables to do the optimization of TAMie parameters within the framework of PC-SAFT which again leads to a significant speed-up [6, 107]. The procedure for optimizing TAMie force field parameters is as follows: First a GCMC simulation is conducted using an initial set of parameters (e.g. from existing force fields or guesses), denoted as \mathbf{p} . A GCMC simulation can be used to generate (p, T, ρ) -equilibrium-data, i.e. saturation pressure p^{sat} at given temperature. The computed properties of interest are denoted by Ω^{sim} . Second the objective function

$$f(\mathbf{p}) = \frac{1}{N^{\text{exp}}} \sum_{n=1}^{N^{\text{exp}}} \left(\frac{\Omega_n^{\text{sim}}(\mathbf{p}) - \Omega_n^{\text{exp}}}{\Omega_n^{\text{exp}}} \right)^2 \quad (2.41)$$

is calculated. With the parameters ϕ_σ and ϕ_ε from equations (C.3) and (C.10) of Appendix C the PC-SAFT model is adjusted to reproduce the simulation data. Using these parameters the PC-SAFT equation is now explicitly expressed in terms of the force field parameters \mathbf{p} (Appendix C) and an auxiliary objective function is formulated as

$$f_{\text{aux}} = \frac{1}{N^{\text{exp}}} \sum_{n=1}^{N^{\text{exp}}} \left(\frac{\Omega_n^{\text{PC-SAFT}}(\mathbf{p}) - \Omega_n^{\text{exp}}}{\Omega_n^{\text{exp}}} \right)^2 \quad (2.42)$$

and is minimized by varying \mathbf{p} . Because this auxiliary function contains only values calculated with the PC-SAFT equation of state, the computational cost of this optimization is very cheap, i.e. taking only a few milliseconds of time. The result of this optimization \mathbf{p} serves as new best guess for the optimal parameters of this force field optimization. Because PC-SAFT only approximates the results of molecular simulations, this procedure is iterative. However, the final converged result is independent from the analytic eos and the procedure converges to the true optimum of the problem [6]. The procedure is repeated until the gradient of f is lower than a specified tolerance.

2.3.4. Determining cross interaction parameters with PC-SAFT

In this section a summary of the determination of group-wise cross interaction parameters κ_{ij} of a molecular force field is given, using the substance-wise cross interaction parameter $k_{\alpha\beta}$ of the PC-SAFT equation of state. The original derivation was done by Schacht et al. [108] and adapted by Waibel et al. [58]. The cross-wise interactions between two van der Waals united-atom Mie interaction sites can be approximated with the Berthelot-Lorentz combining rules [109, 110] for σ_{ij} and ε_{ij} and using an arithmetic mean [111] for

2. Theory and Methods

n_{ij} , as

$$\sigma_{ij} = (\sigma_i + \sigma_j) / 2 \quad (2.43)$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \quad (2.44)$$

$$n_{ij} = (n_i + n_j) / 2. \quad (2.45)$$

The combining rules for σ_{ij} and ε_{ij} are based on London's dispersive theory with identifiable assumptions. The combining rule for the cross-mixed energy parameter ε_{ij} can be allowed to deviate from the Berthelot combining rule, introducing parameter κ_{ij} as

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii} \cdot \varepsilon_{jj}} \cdot (1 - \kappa_{ij}) \quad (2.46)$$

We allow non-zero values for κ_{ij} only for two sites i and j of different type ($i + j$) and site i and site j located on different molecule types. The connection between the cross interaction parameter $k_{\alpha\beta}$ of PC-SAFT and κ_{ij} can be derived by examining a first order perturbation theory on the van der Waals part of the electrostatics of both models and setting them approximately equal

$$\sum_{\alpha} \sum_{\beta} x_{\alpha} x_{\beta} \hat{m}_{\alpha} \hat{m}_{\beta} \hat{\sigma}_{\alpha\beta}^3 \hat{\varepsilon}_{\alpha\beta} \approx \sum_{\alpha} \sum_{\beta} x_{\alpha} x_{\beta} \sum_i^{K_{\alpha}} \sum_j^{K_{\beta}} N_i N_j \sigma_{ij}^3 \varepsilon_{ij}, \quad (2.47)$$

which is equivalent to saying the dimensionless correlation integrals of the PC-SAFT mixture and those of the force field mixture are approximately equal [108]. The indices α and β indicate the two substances of a binary mixture and the index i runs over all types of united-atom groups K_{α} of substance α . N_i marks the number a given UA-group is contained in substance α . Ignoring the pure component entries ($\alpha\alpha$ and $\beta\beta$) of equation (2.47) and focusing on the mixture interactions leads to

$$\hat{m}_{\alpha} \hat{m}_{\beta} \hat{\sigma}_{\alpha\beta}^3 \hat{\varepsilon}_{\alpha\beta} \frac{1}{\varphi_{\alpha\beta}} = \sum_i^{K_{\alpha}} \sum_j^{K_{\beta}} N_i N_j \sigma_{ij}^3 \varepsilon_{ij}. \quad (2.48)$$

The amount-of-substance-fractions $x_{\alpha} x_{\beta}$ have been canceled out and the parameter $\varphi_{\alpha\beta}$ was introduced to enforce the equivalence of the left and right hand side of the equation. Inserting equation (2.46) and analogously for the PC-SAFT equation of state $\varepsilon_{\alpha\beta} = \sqrt{\varepsilon_{\alpha} \varepsilon_{\beta}} \cdot (1 - k_{\alpha\beta})$ finally gives the connection between $k_{\alpha\beta}$ and κ_{ij}

$$k_{\alpha\beta} = 1 - \frac{\sum_i^{K_{\alpha}} \sum_j^{K_{\beta}} N_i N_j \sigma_{ij}^3 \sqrt{\varepsilon_{ij}} (1 - \kappa_{ij})}{\hat{m}_{\alpha} \hat{m}_{\beta} \hat{\sigma}_{\alpha\beta}^3 \sqrt{\hat{\varepsilon}_{\alpha\alpha}} \hat{\varepsilon}_{\beta\beta}} \varphi_{\alpha\beta}. \quad (2.49)$$

2. Theory and Methods

The procedure for optimizing force-field cross-energy parameters κ_{ij} is as follows: First, a GCMC simulation is conducted using $\kappa_{ij} = 0$. A GCMC simulation can be used to generate (p, x^V, x^L) -data, i.e. pressure and mole fractions of coexisting vapor and liquid at given temperature. Using the PC-SAFT model with $k_{\alpha\beta}$ from eq. (2.49), parameter $\varphi_{\alpha\beta}$ is then adjusted such that squared deviations (e.g. in pressure) between results from PC-SAFT and data from molecular simulations are minimal. The PC-SAFT model now approximates results of the force field and parameter $\varphi_{\alpha\beta}$ is considered constant. As a next step the force-field parameter κ_{ij} can be adjusted (using PC-SAFT, still with eq. (2.49)) by minimizing deviations of PC-SAFT to experimental VLE data. Formally the procedure is iterative. In my experience and in agreement with Waibel et al. [58], however, a second iteration is not necessary.

3. TAMie for aldehydes, ketones and small cyclic alkanes

The content of this chapter is a literal quote of the publication

D. Weidler and J. Gross. Transferable anisotropic united-atom force field based on the mie potential for phase equilibria: aldehydes, ketones and small cyclic alkanes. Ind. Eng. Chem. Res., 55(46):12123–12132, 2016

J. Gross had the role of a daily supervisor and was involved in editing the manuscript. Additions or deletions compared to the published work are marked with angular brackets.

The Transferable Anisotropic Mie potential (TAMie) is extended to ketones, aldehydes, and small cyclic alkanes. Parameters defining the Mie pair potential, representing van der Waals interactions, as well as partial point charges were treated as adjustable. These adjustable parameters were identified by minimizing deviation of calculated vapor pressure data and liquid density data to experimental data. The optimization procedure was supported by an analytic equation of state, according to Hemmen and Gross [J. Phys. Chem. B, 2015, 119, 11695–11707]. The optimal point charges for aldehydes and ketones were found by varying them along a predefined grid of values. Results for phase equilibria of pure substances show good agreement with experimental data. The absolute average deviations for aldehydes and ketones in the reduced temperature range $T_R = T/T_{\text{exp}}^{\text{crit}} = 0.55$ to 0.96 are lower than 3.4%. For cyclohexane and cyclopentane the deviations are 0.6%.

3.1. Introduction

Transferable force fields allow predicting the physical properties of substances and mixtures that are weakly characterized by experiments. The optimized potential for liquid simulations (OPLS) is an important transferable force field with emphasis on thermodynamic properties, where force field parameters were adjusted to experimental enthalpies of vaporization and to liquid density data near ambient conditions [44–47]. There are two parameterizations of the OPLS force field. The all-atom (OPLS-AA) model and the

3. TAMie for aldehydes, ketones and small cyclic alkanes

united-atom (OPLS-UA) approach, where a single interaction site is assigned for atoms and the adjoining hydrogen atoms. The intermolecular energy is described with Lennard-Jones pair potentials and point charges, the intramolecular model assumes fixed bond lengths between neighboring interaction sites, so that bond stretching is not taken into account.

A prominent force field for correlating and predicting vapor liquid equilibria is the transferable potential for phase equilibria (TraPPE) which is parameterized for many chemical groups[40, 50–53, 90, 112, 113]. It also assumes Lennard-Jones pair-potentials with united atoms, point charges, and uses fixed bond lengths. TraPPE was developed with emphasis placed on liquid densities ρ and bubble point temperatures. The model gives excellent results for T - ρ projections with good results also for the critical temperature, but it has weaknesses in describing vapor pressures[1, 114].

The NERD force field, impartially named after its developing authors Nath et al. adopts a united-atom model but includes bond stretching[115]. Data for second virial coefficients was considered in the objective function for optimizing the NERD force field. As a consequence this force field leads to a better description of vapor pressures, compared with the TraPPE model. In 2004 it was extended to primary alcohols by Khare et al.[116].

All of these force fields describe the van der Waals interaction of a molecule with LJ-potentials, usually located at the positions of atomic nuclei (with hydrogen taken into account only effectively with the neighboring groups). It was recognized, however, that these force fields are not able to simultaneously describe the vapor pressure and liquid densities over a wide range of temperatures. One approach to overcome this shortcoming to some extent while preserving the Lennard-Jones potential for the van der Waals interaction was proposed by Toxvaerd in 1990. For a CH_3 -group, for example, he introduced an offset between the interaction site representing the united-atom group and the locus of the carbon atom. The offset is meant to account for the hydrogen atoms[117]. This concept is known as the anisotropic united-atom (AUA) model. The offset applied to a CH_2 -group, say in a propane molecule, leads to a branched geometry, with the location of the CH_2 interaction site defined only once the position of all three carbon atoms is defined. For Monte Carlo (MC) simulations, where representative molecular configurations have to be regularly generated, one needs more sophisticated sampling schemes like the configurational biasing scheme by Smit et al. [118]. A comprehensive parameterization of the anisotropic united-atom model (AUA4) was started by Ungerer et al.[119], with several subsequent studies extending the parameterization to various functional groups[120–131]. Our study is concerned with ketones and aldehydes, for which Ferrando et al. proposed AUA force field parameters[126]. The transferability of the force field to mixtures is demonstrated without the necessity to use a binary interaction correction towards Lorentz-Berthelot combining rules.

3. TAMie for aldehydes, ketones and small cyclic alkanes

In 2003 Kranias et al. developed an AUA force field for ketones and aldehydes where they derived atomic charge values with *ab initio* calculations [132]. The approach is appealing because the concept of transferable parameter is thereby only applied to van der Waals potentials, whereas the electrostatic potential is allowed to be individual for all species. A certain drawback lays in the fact that partial charges are thereby assigned from quantum mechanical calculations in vacuum. It is on the one hand rather ambiguous to reduce the single-particle electron density of a molecule to a small collection of point charges. On the other hand, point charges are usually defined as effective non-polarizable point charges that to some extent account for induced polarity.

It is possible to obtain a good description of both, vapor pressure and liquid density when pair potentials other than the Lennard-Jones model are considered. Errington and Panagiotopoulos applied a Buckingham exponential-6 potential and obtained very good results for non-polar substances [114]. Similar to the AUA potential the CH₃-group is shifted outward to effectively account for the hydrogen atoms. Potoff and Bernard-Brunel applied the Mie $n - 6$ potential to describe van der Waals interactions [111]. In treating the repulsive exponents n as an adjustable parameter characterizing a functional group, they found a much improved description of vapor pressure and liquid density when compared to the Lennard-Jones potential (corresponding to a Mie potential with repulsive exponent $n = 12$). Recently this potential was extended to alkenes [133]. Müller and Jackson determine the pure component parameters (Mie potential) of tangent-sphere chain fluids using the analytic SAFT- γ equation of state and use these parameters directly for molecular simulations [55, 56]. This approach has successfully been used for several substance groups, such as greenhouse gases and refrigerants [57], or benzene and n-decylbenzene [134].

Hemmen and Gross also used a Mie $n - 6$ pair-potential for united-atom groups describing the van der Waals contribution to the overall interaction. Like in the AUA force-fields, an offset was allowed for the optimization of the force field parameters, limited, however, to terminal chemical groups, in order to avoid the need for the specific configurational bias scheme mentioned above. For alkanes, for example, six pure force field parameters were simultaneously adjusted to four n-alkanes. Hemmen et al. extended the work towards ethers with a detailed analysis of how point charges are defined [6]. The resulting force field is referred to as Transferable Anisotropic Mie model (TAMie).

In this study we extend the TAMie force field to cyclic alkanes, to ketones, and to aldehydes. The objective function for the optimization problem minimizes deviations in calculated vapor pressure and liquid densities to experimental values.

3. TAMie for aldehydes, ketones and small cyclic alkanes

Table 3.1.: Pair potential parameters of UA-groups: energy parameter ε , size parameter σ , repulsive exponent n , and point charge q for the TAMie force field

pseudoatom i	M_i /(g/mol)	ε_i/k_B /K	TAMie		charge q_i / e	ref.
			$\sigma_i/\text{\AA}$	n_i		
CH ₃ (sp ³)	15.035	136.318	3.6034	14		[1]
CH ₂ (sp ³)	14.027	52.9133	4.0400	14		[1]
O aldehydes	15.999	100.85	3.0276	12	-0.422	this work
CH _{<i>x</i>} neighbor aldehydes					-0.038	this work
CH (sp ²) aldehydes	13.019	68.934	3.4941	12	0.46	this work
O ketones	15.999	65.55	3.093	12	-0.49	this work
C (sp ²) ketones	12.011	32.775	3.919	12	0.49	this work
O acetone	15.999	69.184	3.112	12	-0.49	this work
C (sp ²) acetone	12.011	34.592	3.942	12	0.49	this work
CH ₂ (sp ³) 5-ring	14.027	74.527	3.8462	16		this work
CH ₂ (sp ³) 6-ring	14.027	69.568	3.8967	16		this work

3.2. Force field development

3.2.1. Intermolecular energy

The potential between two interaction sites i and j located on different molecules is described by a Mie potential plus coulombic interactions

$$u_{ij}^{\text{inter}} = c_{ij} \cdot \varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{n_{ij}} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{m_{ij}} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \quad (3.1)$$

where r_{ij} denotes the distance between the interaction sites, σ_{ij} and ε_{ij} are the size parameter and energy parameter characterizing the Mie interaction, and q_i and ε_0 are the point charge and the vacuum permittivity, respectively. An interaction site i can represent a united-atom group, such as a methyl-group CH₃, or a point charge within a united-atom group. The same potential is applied to any pair of interaction sites within a molecule separated by more than three bonds. The constant c_{ij} is chosen to enforce the minimum of the Mie potential at a value of $-\varepsilon_{ij}$. Depending on the choice of repulsive (n_{ij}) or dispersive (m_{ij}) exponent the constant is defined as

$$c_{ij} = \frac{n_{ij}}{n_{ij} - m_{ij}} \cdot \left(\frac{n_{ij}}{m_{ij}} \right)^{\frac{m_{ij}}{n_{ij} - m_{ij}}} \quad (3.2)$$

Throughout this work the dispersive exponent was set to $m_{ij} = 6$ whereas the repulsive exponent n_{ij} was considered a degree of freedom.

We take an advance on the results of our study and list the proposed force field parameters in Table 3.1. Point charges are given as factors of electron charge e . In the TAMie

3. TAMie for aldehydes, ketones and small cyclic alkanes

force field we tabulate parameters describing the interaction of like ($i-i$) pairs, which we denote with a single index (for example as σ_i). The Mie interactions are evaluated in the simulation up to a fixed cutoff radius of $r_c = 14 \text{ \AA}$ and standard tail-corrections were applied. Parameters for unlike interaction sites are determined using the Lorentz-Berthelot combining rules [109, 110] for σ_{ij} and ε_{ij} and using an arithmetic mean[111] for n_{ij} , as

$$\sigma_{ij} = (\sigma_i + \sigma_j) / 2 \quad (3.3)$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \quad (3.4)$$

$$n_{ij} = (n_i + n_j) / 2 \quad (3.5)$$

The combining rules for σ_{ij} and ε_{ij} are based on London’s dispersive theory with transparent assumptions. The combining rule for the repulsive exponent n_{ij} is more ambiguous. Stiegler and Sadus evaluated how various choices of defining cross-potentials affect the physical properties of the resulting mixture [135]. With the objective of correlating and predicting real mixtures, one could draw justification for a certain combining rule from quantum mechanical calculations. For transferable force fields, the choice of one over the other combining rule is pragmatically justified by favorable predictions of experimental mixture data.

3.2.2. Intramolecular energy

The proposed force field parameterization uses fixed bond lengths between interaction sites. The intramolecular energy has three contributions: the bending energy, the torsional energy and the site-site energy, eq. (5.1), for pairs separated by more than three bonds. We adopt intramolecular parameterizations from the TraPPE force field [50, 53, 113] which is partly based on parameterizations of the OPLS force field [44], the AMBER force field [19] and force constants proposed by van der Ploeg and Berendsen[136]. More specifically, the angle potentials stem from the AMBER force field [19]. The torsion potentials are based on OPLS[44, 49]. Exceptions are cyclic alkanes where an original intramolecular parameterization was proposed for TraPPE[50, 113].

The angle bending potential is represented as a simple harmonic potential

$$u_{\text{bend}} = \frac{k_0}{2} (\theta - \theta_{eq})^2 \quad (3.6)$$

where θ_{eq} is the temperature-independent equilibrium angle and k_0 is the force constant. For cyclic molecules the torsional potential is defined as

$$u_{\text{torsion}}^{\text{cyc}} = c_0 + c_1 \cos(\phi) + c_2 \cos(2\phi) + c_3 \cos(3\phi) \quad (3.7)$$

3. TAMie for aldehydes, ketones and small cyclic alkanes

Table 3.2.: Bond lengths used in the proposed TAMie force field taken from literature.

type	$r_0 / \text{\AA}$	ref
$\text{CH}_3\text{-CH}_x \quad x \in \{1, 2\}$	$1.54 + 0.2$	[1, 6]
$\text{CH}_x\text{-CH}_y \quad x, y \in \{1, 2\}$	1.54	[50]
$\text{CH}_3\text{-CH[=O}_{\text{ald}}]$	$1.52 + 0.2$	
$\text{CH}_2\text{-CH[=O}_{\text{ald}}]$	1.52	[53]
$\text{CH}_3\text{-C[=O}_{\text{ket}}]$	$1.52 + 0.2$	
$\text{CH}_2\text{-C[=O}_{\text{ket}}]$	1.52	[53]
$\text{CH}_2\text{-CH}_2$ (cyclic)	1.54	

Table 3.3.: Bending angles and constants for the proposed TAMie force field taken from literature.

bending site	$\theta / ^\circ$	$k_0/k_B / \text{K/rad}^2$	ref
$\text{CH}_x\text{-CH}_2\text{-CH}_y$	114.0	62 500	[50, 136]
$\text{CH}_x\text{-C}_{\text{ket}}\text{-CH}_y$	117.2	62 500	[53]
$\text{C}_{\text{ket}}\text{-CH}_2\text{-CH}_x$	114.0	62 500	[53]
$\text{CH}_x\text{-C}_{\text{ket}}\text{=O}_{\text{ket}}$	121.4	62 500	[53]
$\text{CH}_x\text{-CH=O}_{\text{ald}}$	121.4	62 500	[53]
$\text{CH}_2\text{-CH}_2\text{-CH}_2$ (cyc5)	105.5	62 500	[113]
$\text{CH}_2\text{-CH}_2\text{-CH}_2$ (cyc6)	114.0	62 500	[50]

while for all other substances considered in this study, it is defined as

$$u_{\text{torsion}} = c_0 + c_1 [1 + \cos(\phi)] + c_2 [1 - \cos(2\phi)] + c_3 [1 + \cos(3\phi)] \quad (3.8)$$

Bond lengths, parameters of angle bending potentials, and parameters of torsional potentials are summarized in Table 3.2 to 3.4, respectively.

Table 3.4.: Torsional potential constants taken from literature.

torsion sites	$c_0/k_B / \text{K}$	$c_1/k_B / \text{K}$	$c_2/k_B / \text{K}$	$c_3/k_B / \text{K}$	ref
$\text{CH}_x\text{-CH}_2\text{-CH}_2\text{-CH}_y$	0.0	355.03	-68.19	791.32	[44, 50]
$\text{CH}_x\text{-CH}_2\text{-C=O}_{\text{ket}}$	2035.58	-736.90	57.84	-293.23	[53]
$\text{CH}_x\text{-C-CH}_2\text{-CH}_y$	-17.26	752.60	14.89	282.10	[54]
$\text{C-CH}_2\text{-CH}_2\text{-CH}_x$	0.0	355.03	-68.19	791.32	[44, 50]
$\text{CH}_x\text{-CH}_2\text{-CH}_2\text{-CH[=O}_{\text{ald}}]$	11.81	467.80	-274.10	846.80	[53]
$\text{C-CH}_2\text{-CH=O}_{\text{ald}}$	1182.0	-225.60	302.80	-339.30	[54]
$\text{CH}_2\text{-CH}_2\text{-CH}_2\text{-CH}_2$ (cyc5)	31394	45914	16518	1496	[113]
$\text{CH}_2\text{-CH}_2\text{-CH}_2\text{-CH}_2$ (cyc6)	5073	6840	3509	63	[113]

3.2.3. Optimization of force field parameters

The effective two-body van-der Waals and the effective (non-polarizable) point charge parameters were treated as adjustable parameters. Parameter optimization was performed using the method described by van Westen et al.[107] and modified by Hemmen and Gross[1], where an analytical equation of state (Perturbed-Chain Statistical Associating Fluid Theory, PC-SAFT equation of state[59, 60]) is used to approximate the objective function locally. The optimization procedure is iterative, because first, the equation of state gives approximate results and more importantly, second, the molecular model underlying the PC-SAFT equation of state is more coarse compared to the TAMie force field. The convergence, however, is very swift. Usually only three to five iterations are needed for a sufficiently converged result. Hemmen and Gross modified the optimization procedure slightly and showed, that approximations of the equation of state (in the optimum) do not act on the optimum of the force-field optimization problem. The algorithm thus converges to the true minimum of the problem of force-field optimization; the analytical model does not alter the converged result.

Point charges and repulsive exponents, however, were not optimized simultaneously with parameters of the Mie potential (ε_i and σ_i). Rather, we vary point charges along a predefined grid of values and define the optimal point charges *a posteriori*. The optimum of the Mie energy parameter ε_i and the size parameter σ_i was reached in an automated process, where molecular simulations and intermediate PC-SAFT optimization steps are successively executed.

The objective function to be minimized is defined as a sum of squared relative deviations of simulated ($\Omega_n^{\text{sim}}(\mathbf{p})$) to experimental observables, $\Omega_n^{\text{exp}}(\mathbf{p})$, as

$$f(\mathbf{p}) = \frac{1}{N^{\text{exp}}} \sum_{n=1}^{N^{\text{exp}}} \left(\frac{\Omega_n^{\text{sim}}(\mathbf{p}) - \Omega_n^{\text{exp}}}{\Omega_n^{\text{exp}}} \right)^2 \quad (3.9)$$

where \mathbf{p} is the set of adjustable force-field parameters, N_{exp} is the number of quasi experimental data points. In our objective function, the observables Ω include liquid density ρ_{liq} data and vapor pressure p^{sat} points. Quasi-experimental data was taken from multi-parameter correlations of the DIPPR data base [137], the Korean Thermophysical Properties Database (KDB)[138] and the Dortmund Database (DDBST)[139].

3.2.4. Simulation details

All simulations were performed in grand canonical ensemble, using a histogram reweighting technique in a post-processing step. A detailed description of the simulation technique is given in previous work[1, 6] and we proceed with a short overview of these methods.

In grand canonical ensemble the number of molecules N fluctuates. We used the tran-

3. TAMie for aldehydes, ketones and small cyclic alkanes

sition matrix method [78, 140] for determining a bias potential on the fly, ensuring an approximately equal sampling of molecule number N . Additionally we split the N -space, i.e. the coordinate of varying molecule number, into different windows of size ΔN . All windows can be combined to a total histogram by applying the multiple ensemble technique described by Lyubartsev et al. [141]. The decomposition into many simulations covering a certain domain of N to $N + \Delta N$ has the advantage, that all windows can run in parallel. A message passing interface (MPI) implementation automated the process of generating starting values, splitting the particle space in different windows, performing histogram reweighting, and calculating new optimized parameters. The number of windows and thus the width ΔN of each window was chosen so that the total number of threads could be sent to a compute cluster. The width of a window was approximately $\Delta N = 10$ molecules. Virnau and Müller showed that the statistical uncertainties does not depend on the number of windows into which the N -domain is divided [142].

The volume of the simulation box varied between 40000 Å for propanal and 81000 Å for octanal. Simulations were carried out with at least 5 million equilibration steps and 25 million production MC-steps for each window of size ΔN . The probabilities of choosing different moves were set to 60 % for translation/rotation, 30 % for insertion/deletion and 10 % for regrowing a molecule for all substances. Configurational bias (CB) ensures sufficiently high probabilities of inserting or deleting molecules at higher densities. The number of CB trial-steps was set to 1 for low numbers of molecules and increasing towards 8 for the highest densities.

For branched molecular structures, such as ketones and for cyclic alkanes it is necessary to obtain intramolecular trial configurations iteratively. We apply a MC scheme with importance sampling on isolated molecules to develop representative trial configurations. Trial structures are needed for growing a new molecule, for regrowing an existing molecule, or for determining the Rosenbluth weight when trial-deleting a molecule. The generation of these configurations occurs during runtime up to a maximum number $N^{\text{trial-structures}}$.

For cyclic molecules we apply the MC scheme to iteratively determine the configuration of the entire molecule, whereas for branched molecules, the branching node with neighboring interaction sites is iteratively obtained. All trial configurations of these structures (full cyclic molecule or branching node with neighboring sites) are stored in a library until a maximum number of $N^{\text{trial-structures}}$ structures is reached. Subsequently, the configuration of a random structure is drawn from the library whenever a trial configuration is needed. This procedure leads to a significantly improved computational efficiency compared with an incessant generation of new structures. Similar approaches were proposed in refs. [87, 89, 114]. The library of molecular configurations is a simple array in random-access memory.

For the cyclic molecules the molecular structures are generated through crank-shaft MC

moves [91, 92] that change the bending and torsional angles. We use 2000 trial moves and accept the crank-shaft move with an importance-sampling acceptance probability. The first $N^{\text{trial-structures}} = 5 \cdot 10^5$ so-determined molecular structures are stored in the library.

Cyclic molecules are here composed of three CB groups: The first interaction site, the second site, and all other sites of the molecule. The first interaction site is inserted following the normal configurational biasing scheme (using 8 random trial positions for the highest densities). The second interaction site, is similarly placed randomly at a bond-length around the first site. Subsequently, a cyclic structure is newly built or drawn from the library. The so-defined bond-angle between first, second, and third interaction site allows for trial positions of the third interaction site located on an arc. All other interaction sites are then defined.

For coulombic potentials we apply the standard Ewald summation. The damping parameter α of the Ewald summation was set to $\alpha = 7.59$ with the maximum number of k -vectors $k_{max} = 9$ in each direction.

Phase equilibria were calculated in a post-processing step with the histogram reweighting technique as proposed by Ferrenberg and Swendsen [72, 73] and later adopted for Lennard-Jones fluids by Wilding[74] and for polymers by Panagiotopoulos et al.[75]. Details of the implementation of histogram reweighting is given in the study of Hemmen et al.[6]. Mixed-field finite-size scaling techniques [74, 75] and critical scaling laws were applied to determine the critical point.

3.3. Results

3.3.1. Small cyclic alkanes

We assign individual parameters for the 5-ring and the 6-ring of cyclopentane and cyclohexane and optimized σ_{CH} and ϵ_{CH} as well as the repulsive exponent n_{CH_2} of each of the cyclic CH-groups. No point charges are used. Parameters defining the intramolecular potential are given in Table 3.2 to 3.4. We analyzed the molecular 'boat' and 'chair' conformations using the criterion proposed by Keasler et al.[143]. For cyclohexane in the vapor phase at $T = 450$ K we found the probabilities $f_{\text{chair}} = 0.994$ and $f_{\text{boat}} = 0.005$, and for a liquid phase at $T = 510$ K $f_{\text{chair}} = 0.989$ and $f_{\text{boat}} = 0.009$, respectively. With a small probability the molecular structures could not be classified as either of the two conformations. Several earlier studies have optimized force field parameters to experimental data [114, 120, 144, 145]. The results of their parameter optimization is different to our work, because different degrees of freedom are considered. We treat the repulsive exponent of the Mie potential as a degree of freedom (constrained, however, to even integer values), whereas other studies constrain the exponent to a value of $n_{\text{CH}_2} = 12$. The

3. TAMie for aldehydes, ketones and small cyclic alkanes

degree of freedom is comparable to the degree of freedom selected in the study of Errington and Panagiotopoulos for the Buckingham potential[114]. The study of Muñoz-Muñoz et al.[144] on cyclic alkanes is conceptually interesting, because their objective function included both, static properties and transport properties.

For both, cyclopentane and cyclohexane, the optimal repulsive exponent is $n_{\text{CH}_2} = 16$ (Table 3.1). Because the optimization of van der Waals parameters is done individually, we obtain very good agreement to experimental data included in the objective function. Figure 3.1 shows the saturation pressures of small cyclic alkanes calculated with the TAMie force field. (The figure shows the logarithm of the pressure versus the inverse temperature which gives a almost linear representation of the data. To make it easier for the reader the axis are labeled with the absolute values of the vapour pressure p^{sat} and the temperature T and the scale of the axis was adjusted accordingly. All following figures of that type are presented this way throughout this thesis.) For comparison the vapor pressures calculated with the TraPPE force fields are included. Figure 3.3 allows a refined assessment of the deviations between calculated and experimental data. The average errors (AAD) in vapor pressure (for the temperature range from $T_R = T/T_{\text{exp}}^{\text{crit.}} = 0.55$ to 0.96) are 0.4% and 0.4% for cyclohexane and cyclopentane, respectively. Average deviations in liquid densities for both substances are 0.6% and 0.5%, respectively. Critical point properties are not part of our objective function, which is why we didn't analyze the critical point in detail. The critical points are 0.6% and 1.3%, respectively, higher than the experimental values. Tabulated simulation results (including enthalpies of vaporization) are reported in the supporting information.

3.3.2. Ketones

The physical properties of the smallest member of a chemical family are usually not captured well in transferable force fields. Like in our previous works on the TAMie force field we excluded the smallest member of the homologous series (here acetone). If we include acetone in our objective function we find a markedly higher overall error in describing the chemical family. This observation either reveals limits to the concept of transferable parameters, or limits of the simplistic molecular model, assuming constant (non-polarizable) point charges and omitting hydrogen atoms. We consider acetone individually and optimize Mie potential parameters of the acetone-oxygen and of the acetone-carbon for varying partial charges. We use constant effective point charges located, for ketones, at the interaction sites of the Mie groups. The OPLS force field assigns point charges for the ketone-oxygen and on the ketone-carbon and small *positive* partial charges on the neighboring methyl- groups. Kamath et al. performed a quantum mechanical investigation on the point charges of acetone by considering a cluster of acetone with a chloroform molecule[146]. They identify mildly *negative* point charges for the methyl-groups

3. TAMie for aldehydes, ketones and small cyclic alkanes

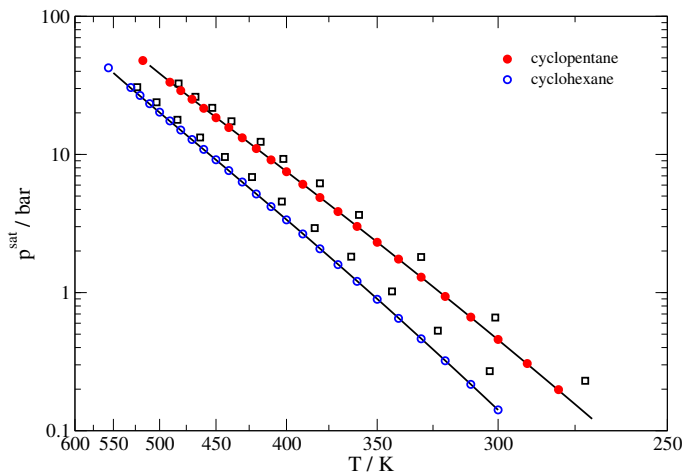


Figure 3.1.: Vapor pressure of cyclopentane and cyclohexane. Comparison of results from the TAMie force field (spheres) with quasi-experimental data[137] (lines). Results of the TraPPE force field [143] (open squares) are also included.

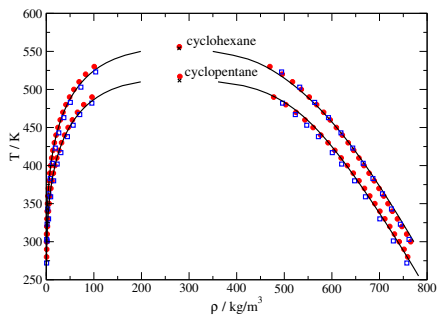


Figure 3.2.: Vapor-liquid coexisting densities for cyclic alkanes. Comparison of results from the TAMie force field (filled spheres) with quasi-experimental data[137] (lines). Results of the TraPPE force field [143] (open squares) are also included.

3. TAMie for aldehydes, ketones and small cyclic alkanes

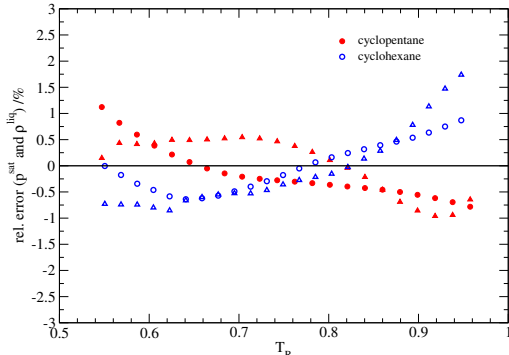


Figure 3.3.: Relative errors of the TAMie force field from quasi-experimental data for cyclopentane (blue empty symbols) and cyclohexane (red solid symbols) for varying reduced temperature $T_R = T/T_{\text{exp}}^{\text{crit.}}$. Deviations in the vapor pressure (circles) and in liquid density (squares) are included.

of acetone (in contrast to the OPLS parameterization). The TraPPE force considers the methyl-groups as electro-neutral. Similarly, we decided to assign point charges only to oxygen and the carbon atom of the ketone-group, with $q_{\text{O}(\text{ketone})} = -q_{\text{C}(\text{ketone})}$, and treat the neighboring groups as neutral.

The Mie potential parameters of the oxygen and the carbon atom within the ketone-group are highly correlated. We therefore defined the ratios of $\epsilon_{\text{O}(\text{ketone})}/\epsilon_{\text{C}(\text{ketone})}$ and $\sigma_{\text{O}(\text{ketone})}/\sigma_{\text{C}(\text{ketone})}$ as fixed and equal to the ratios of the OPLS force field. We thus have three degrees of freedom, i.e. $\epsilon_{\text{O}(\text{ketone})}$, $\sigma_{\text{O}(\text{ketone})}$, and $q_{\text{O}(\text{ketone})}$, for optimizing the force field parameters of acetone to experimental data. An initial investigation showed no advantage in using repulsive Mie-exponents for the ketone-carbon and -oxygen group other than $n = 12$. The result of the parameter optimization is shown in Fig. 3.4 for varying partial charges, $q_{\text{O}(\text{acetone})}$. Each point in this diagram minimizes the objective function with respect to the other two parameters, $\epsilon_{\text{O}(\text{acetone})}$ and $\sigma_{\text{O}(\text{acetone})}$. It is interesting to note, that the optimum is found for point charges that almost exactly reproduce the vacuum value dipole moment of acetone[137]. That is unexpected, because the static polarizability of molecules leads to higher effective dipole moments for dense states. It is further interesting to note, that in the optimum, the absolute average deviation (AAD) in liquid density is rather high, with 1.93%, the error in vapor pressure is very low, with 0.78% and the critical point is overestimated with 1.4%. Fig. 3.4 shows the two contributions to the overall error, i.e. the contribution from vapor pressure and from liquid densities. Although not part of the objective function we also include errors of critical temperature. Increasing partial (point) charges lowers the errors in liquid density (and critical temperature), whereas errors in vapor pressure start to steeply increase at

3. TAMie for aldehydes, ketones and small cyclic alkanes

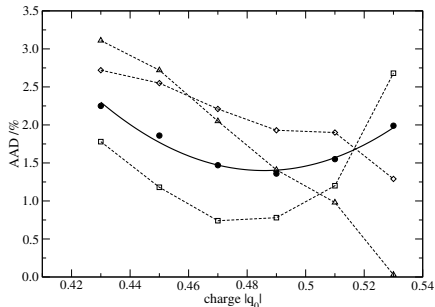


Figure 3.4.: Acetone: Absolute average deviations (%) of simulated vapor pressure and liquid density data to experimental data for varying partial charge q_O of the oxygen (solid spheres). The temperature range corresponds to $0.53 \leq T/T_{\text{exp}}^{\text{crit.}} \leq 0.97$. For each symbol (solid sphere), the two parameters $\epsilon_{O(\text{acetone})}$ and $\sigma_{O(\text{acetone})}$ (with $n_O = 12$) were optimized. The open symbols show the error contributions: vapor pressure (squares), liquid density (diamonds), and critical point (triangles, not part of our objective function).

some point. The optimum thus represents a compromise of errors in liquid density and in vapor pressure.

We now consider the simultaneous adjustment of the parameters $\epsilon_{O(\text{ketone})}$ and $\sigma_{O(\text{ketone})}$ to butanone, 2-pentanone, and 2-octanone. For ketones, we had a similar observation, as earlier reported for ethers[6]: partial charges individually optimized for butanone and 2-pentanone are approximately $q_{O(\text{ketone})} = -0.49$, i.e. in the vicinity of the value identified for acetone. When force field parameters are adjusted simultaneously to several ketones while treating the partial charge as a degree of freedom, however, we find optimized values for partial charges significantly different (here: higher). We decided to prescribe a value of the partial charge $q_{O(\text{ketone})} = -0.49$, because we expect advantages in predicting mixture properties. The optimization with respect to the parameters $\epsilon_{O(\text{ketone})}$ and $\sigma_{O(\text{ketone})}$ leads to an AAD value of 4.0%. The average deviation is dominated by errors in vapor pressure, whereby the vapor pressure is calculated at too high values for butanone and at too low values for 2-pentanone and 2-octanone. Table 3.5 summarizes deviations in vapor pressure, liquid density and critical points. The supporting information provides tabulated simulation results. For 2-pentanone and 2-octanone a comparison to results of the TraPPE force field is also included.

3.3.3. Aldehydes

A TAMie parameterization for aldehydes requires point charges as well as parameters of the Mie potential for the aldehyde-oxygen and the aldehyde-carbon atom. We do not attempt to include methanal (formaldehyde) as a member of the aldehyde family with

3. TAMie for aldehydes, ketones and small cyclic alkanes

Table 3.5: Absolute average deviations (liquid density, saturation pressure) of the TAMie force field from quasi-experimental data[137, 138]. Averages are for reduced temperatures from $T_R = 0.55$ to 0.97 in intervals of $\Delta T = 10$ K.

substance	AAD p^{sat} /%	AAD ρ_l /%	error T_c /%
acetone	0.78	1.93	2.31
butanone	9.39	2.59	0.54
2-pentanone	3.41	1.33	1.51
2-octanone	6.14	1.78	1.87
propanal	5.41	1.29	0.87
butanal	3.12	0.79	0.96
pentanal	5.70	1.20	1.58

transferable force field parameters. Due to the different charge distribution for methanal compared with other aldehydes, we view methanal as an own species. For ethanal, we observed significant scatter and uncertainty in the experimental data, in particular in vapor pressure data. We therefore excluded ethanal in this investigation. Similar to the procedure described in the previous section (for ketones), we started optimizing the force field parameters of propanal. As part of their work in developing the TraPPE force field, Stubbs et al. observed that the original OPLS partial charges were not suited to reproduce phase equilibrium properties [53]. They proceeded with a quantum mechanical study (MP2, with 6-31G(d,p) basis set) adjusting partial charges. We adopted the charge ratios between the aldehyde-oxygen (O -group), the aldehyde carbon (CH -group), and the neighboring carbon-group (CH_x -group) from their work. As a degree of freedom, we linearly scaled the charges proposed by Stubbs et al. As further degrees of freedom, we consider the parameters of the Mie potential for the aldehyde CH -group and the aldehyde O -group. We kept the ratio $\sigma_{CH}/\sigma_O = 1.154$ and the ratio $\epsilon_{CH}/\epsilon_O = 1.463$ fixed during the parameter optimization in order to eliminate highly correlated degrees of freedom. The OPLS-UA force field would have been a natural source for the parameter ratios, because the intramolecular force field of TAMie (like for TraPPE) is to a large extent based on OPLS. However, a OPLS-UA force field for an aldehyde group does not exist, to our knowledge. We therefore took these ratios from the TraPPE force field [53]. Optimization of the force field parameters for propanal showed lowest deviations for a charge of $q_{(CH)aldehyde} = 0.46$ with average errors of vapor pressure and liquid density of about 1%, as Fig. 3.5 shows. The figure also gives results obtained when Mie parameters are simultaneously optimized to propanal, butanal, and pentanal for varying partial charge. Similar to the ketones (and to ethers[6]) an unconstrained simultaneous optimization of Mie parameters and partial charges to several members of a chemical family (propanal, butanal, and pentanal) gives partial charge values rather different from values adjusted individually. We reapply our earlier approach, according to which the partial charges are

3. TAMie for aldehydes, ketones and small cyclic alkanes

taken from one representative member of a homologous series - propanal in this case. The Mie parameters, however, are determined from a simultaneous optimization. The open symbol at $q_{(\text{CH})\text{aldehyde}} = 0.46$ in Fig. 3.5 thus represents the proposed TAMie parameterization.

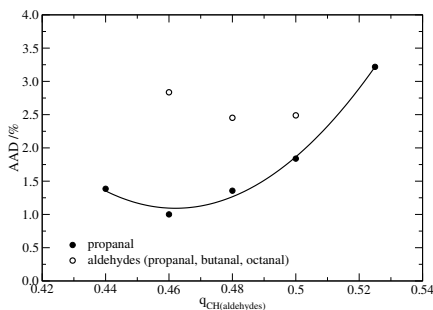


Figure 3.5.: Aldehydes: Absolute average deviations (%) of simulated vapor pressure and liquid density data to quasi-experimental data[138] for varying partial charge q_{CH} of the aldehyde carbon CH -group. The temperature range covers $0.5 \leq T/T_{\text{exp}}^{\text{crit.}} \leq 0.97$. For each symbol, the Mie parameters of the aldehyde group were optimized. For the open symbols, the optimization was conducted simultaneously for propanal, butanal, and pentanal.

Fig. 3.6 and 3.7 compare results obtained for the TAMie force field to experimental data for coexisting densities and vapor pressures, respectively. Both diagrams confirm good agreement to the experimental data of propanal, butanal and pentanal, as members of the objective function. Fig. 3.7 also reveals significant scatter in experimental data of vapor pressures for this chemical family.

3.4. Transferability of parameters

To assess the transferability of the proposed force field we consider substances of the same homologous series that were not part of the objective function during parameter optimization. For aldehydes we consider hexanal and octanal as components that were not included in adjusting the force field parameters. The TAMie-results for these substances are included in Fig. 3.6 and 3.7 and they are found in good agreement to the experimental data.

We now consider ketones. The objective function only included 2-ketones, so that the comparison of TAMie predictions for 3-pentanone to experimental data is an important

3. TAMie for aldehydes, ketones and small cyclic alkanes

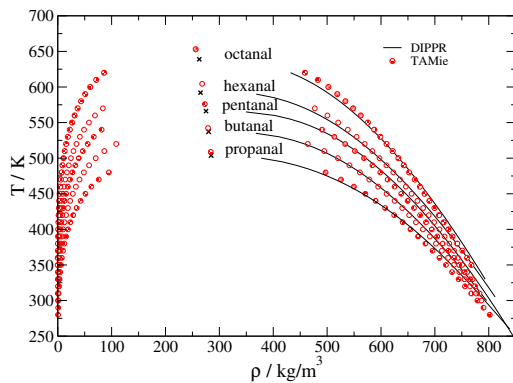


Figure 3.6.: Phase diagrams of simulated aldehydes. Lines are quasi-experimental data[137]. Symbols represent simulation results using the TAMie force field.

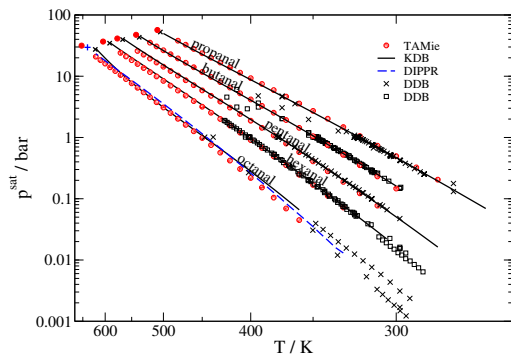


Figure 3.7.: Vapor pressures of aldehydes: crosses and squares represent experimental data, the lines are correlations of KDB[138], and DIPPR[137], respectively. Red spheres represent simulated values.

3. TAMie for aldehydes, ketones and small cyclic alkanes

Table 3.6.: Absolute average deviations of TAMie-predictions from quasi-experimental data[137] for vapor pressure, liquid density, and critical temperature of two ketones not considered during parameter optimization. Averages are for the temperature range $T_R = 0.55$ to 0.97

Substance	AAD p^{sat} /%	AAD ρ_l /%	error T_c /%
3-pentanone	1.62	3.40	0.16
2-hexanone	2.90	1.84	1.6

test. Table 3.6 shows the deviations for ketones that were not considered for the parameter optimization. For 3-pentanone and 2-hexanone we obtain average deviation in vapor pressure of about 2.2%. The errors in liquid density of 2.6% are higher than expected, because liquid density data is usually comparably easy to predict. The predictions of the critical temperature, however, are good, as Table 3.6 confirms. For 3-pentanone we also performed molecular simulations using the TraPPE force field[53]. For TraPPE we obtain AAD values of 1.5% and 29.6% for liquid densities and vapor pressures, respectively. The comparison to TAMie (Table 3.6) confirms markedly improved results for vapor pressure data with some sacrifice on liquid densities.

Fig. 3.8 shows the deviations of the predictions from TAMie from quasi-experimental data for both ketones. The diagram gives deviation of vapor pressure as well as in liquid densities.

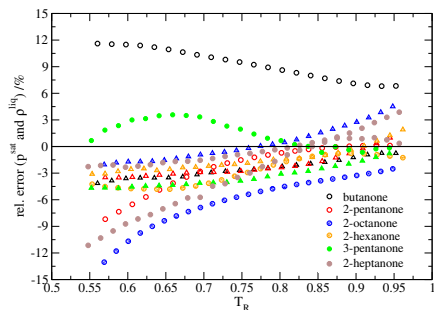


Figure 3.8.: Relative errors of ketones compared to quasi-experimental correlations[137]. Triangles represent errors of the liquid density and circles the errors of the vapor pressure.

4. Individualized force fields based on the transferable anisotropic Mie potential

The content of this chapter is a literal quote of the publication

D. Weidler and J. Gross. Individualized force fields for alkanes, olefins, ethers and ketones based on the transferable anisotropic mie potential. Fluid Phase Equilib., 470:102–108, 2018

J. Gross had the role of a daily supervisor and was involved in editing the manuscript. Additions or deletions compared to the published work are marked with angular brackets.

Transferable force fields allow the prediction of physical properties for substances with scarce or absent experimental data. For substances with a comprehensive experimental database, however, transferable force fields produce results with higher errors than desired, because the transferable force fields are designed to represent a compromise in correlating the properties of many substances. For applications in chemical engineering, where requirements for accurate vapor pressure correlations and predictions exist, the results from transferable force fields are sometimes insufficient. We individualize the transferable anisotropic Mie force field (TAMie) for 38 substances of various chemical families, by introducing a correction parameter that scales all van der Waals energy parameters ε_i of the considered substance. We find markedly reduced deviations, mainly in the description of vapor pressures while the errors in liquid density do not change significantly. For polar species, the improvement in vapor pressure is typically a factor four in absolute average deviation to experimental data, when compared to the original TAMie force field. The improved description of pure substances enables more reliable predictions of phase equilibria in binary mixtures. The concept of individualizing substances that are well-characterized by experimental data within a transferable force field is appealing, because the ability to predict mixtures with substances that are not covered by experimental data is preserved.

4.1. Introduction

Transferable force fields allow predicting physical properties and phase equilibria of pure substances and mixtures that are poorly described by experimental data. Among the most established transferable force fields are the Transferable Potentials for Phase Equilibria (TraPPE) [40, 50–53, 90, 112], the Optimized Potential for Liquid Simulations (OPLS) [44–49] and the Assisted Model Building with Energy Refinement (AMBER) force field [19].

The AMBER force field as well as the TraPPE-UA force field are based on so-called united-atom models, where the van der Waals contribution of hydrogen atoms are lumped with interaction sites of neighboring atoms. The van der Waals interaction of a methyl-group ($-\text{CH}_3$), for example, is considered as one effective interaction site. Recent developments improved the description of properties like the vapor pressure by introducing additional degrees of freedom in the parameterization of transferable force fields. An additional degree of freedom can be introduced for united-atom approaches, by moving the interaction site of, say a CH_3 group, away from the position of the carbon atom to better (effectively) account for the presence of hydrogen atoms. This approach was proposed by Toxvaerd[117, 147] and led to the Anisotropic United Atom force field [123–131, 148] (AUA) by Ferrando, Boutin, Ungerer and coworkers. Shifting of interaction sites of non-terminal groups (say a CH_2 -group) outwards requires more advanced configurational bias sampling schemes, as proposed by Smit et al.[118], because the location of the CH_2 -interaction site is only defined once both neighboring interaction sites have a defined location. The method of Smit et al. introduces the small offset of the interaction sites as a last step in a configurational bias sampling scheme.

Other studies use intermolecular potentials other than the Lennard-Jones potential, e.g. the Buckingham potential with three adjustable parameters per van der Waals interaction site[114] or the Mie potential also with three adjustable parameters [111, 133, 149, 150]. These works show that a united atom force field build up on potentials other than Lennard-Jones can provide good results for vapor pressures and densities for phase coexistence.

The Transferable Anisotropic Mie (TAMie) force field, recently developed in our group with emphasis on thermodynamic properties and vapor-liquid equilibria[1, 3, 6] adopts some of these ideas. It is a united-atom model and incorporates anisotropy in a simple manner by allowing an offset to terminal united-atom sites. With this simple approach sophisticated sampling schemes like the one proposed by Smit et al.[118] can be avoided. A Mie potential is used for the van der Waals contribution of inter- and non-covalent intramolecular interactions.

For the multidimensional optimization of the force field parameters an analytic equation of state, namely the Perturbed Chain Statistical Associating Fluid Theory (PC-SAFT)

4. Individualized force fields based on the transferable anisotropic Mie potential

equation of state[59, 60] was used. Although the molecular model of the PC-SAFT model is more coarse compared to the TAMie force field, it is possible to approximately map the objective function of TAMie to an analogous objective function of PC-SAFT. The procedure is iterative but converges rapidly[107] and the final converged force field parameters are not biased by the PC-SAFT model[1].

Another approach of coupling an equation of state with molecular simulations was proposed by Ghobadi and Elliott who applied a SAFT with molecular parameters taken from force fields that were originally proposed for molecular simulations [151–153]. Recently, Müller and Jackson with coworkers derived force field parameters directly from a physically based equation of state (the SAFT- γ equation of state)[55]. Their procedure is appealing because it is non-iterative and leads to good results for thermodynamic properties and phase equilibria. The molecular model is limited, of course, to the molecular model underlying the analytic equation of state. Their molecular model considers molecules as tangentially bonded Mie interaction sites without bond-angle potentials. The usefulness of analytic equations of state for conducting molecular simulations has in another context been demonstrated by Gospodinov and Escobedo, who showed how probability distributions can be estimated from equations of state[154, 155]. Sanchez et al. [156] used the PC-SAFT model for estimating bias potentials that allow sampling across the energetic barrier between a vapor and a liquid phase in open ensembles.

Substances that are comprehensively characterized by experimental data are less accurately described with a transferable force field compared to a force field individually adjusted to data of that substance. Deficiencies in describing individual substances are a result of a compromise that needs to be made in correlating force field parameters to many substances simultaneously. For the TAMie force field this compromise leads to a positive deviation in the vapor pressure curve of some substances and a negative deviation from the experimental vapor pressure others, especially for polar substances. For chemical engineering applications, deviations in vapor pressure are critical for process safety and equipment design aspects and should be eliminated as much as possible. Figure 4.1 visualizes the systematic offset in vapor pressure for ethers (di-ethyl-ether and di-n-propyl-ether) as calculated from the TAMie force field. The diagram shows relative deviations of calculated vapor pressure and the liquid density from experimental data for both substances. One observes overpredicted vapor pressure values for di-ethyl-ether and underpredicted values for di-n-propyl-ether, when using the optimum for the given degrees of freedom[6].

In this study we propose the concept of individualized transferable force fields. The concept preserves the strength of a transferable force field in predicting properties of substances that are weakly characterized by experimental data, while alleviating the problem of systematic deviations (mainly in vapor pressure) found for some species. We introduce

4. Individualized force fields based on the transferable anisotropic Mie potential

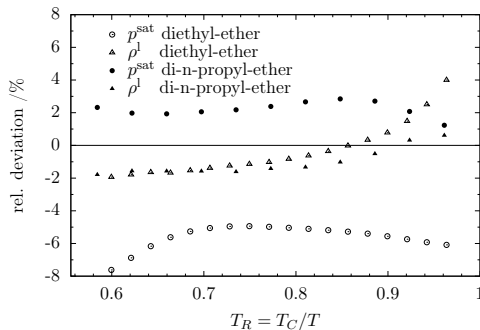


Figure 4.1.: Relative errors of the TAMie force field for ethers: spheres denote relative deviations concerning the vapor pressure, while the triangles symbolize the relative deviations of the liquid density. Filled symbols belong to dipropyl-ether and empty symbols to diethyl-ether.

a correction parameter ψ_A that scales all van der Waals energy parameters of a pure substance A and shifts the vapor pressure curve closer to experimental data. We present individualized force field parameters for 38 substances and show that mixture properties are better described.

4.2. Methodology

4.2.1. The TAMie force field

The TAMie force field is currently parameterized for n-alkanes and n-olefins[1], ethers[6], small cyclic alkanes, ketones, and aldehydes[3]. The pair potential between two interaction sites located either on different molecules, or more than three bonds apart in the same molecule writes as the sum of a Mie potential and an electrostatic potential, as

$$u_{ij} = c_n \cdot \varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{n_{ij}} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi \varepsilon_0 r_{ij}} \quad (4.1)$$

where r_{ij} is the distance between two interaction sites i and j , ε_{ij} is the van der Waals energy parameter, σ_{ij} is the corresponding size parameter, n_{ij} denotes the repulsive exponent, and q_i is the partial point charge. The partial charge or the energy parameter can be zero. Furthermore, ε_0 is the vacuum permittivity and the constant c_n ensures the

4. Individualized force fields based on the transferable anisotropic Mie potential

minimum of the Mie potential at $-\varepsilon_{ij}$, with

$$c_n = \left(\frac{n_{ij}}{n_{ij} - 6} \right) \left(\frac{n_{ij}}{6} \right)^{\frac{6}{n_{ij}-6}} \quad (4.2)$$

The cross-wise parameters of the Mie potential are calculated using the Lorentz-Berthelot combining rules[109, 110]

$$\sigma_{ij} = \frac{1}{2} (\sigma_{ii} + \sigma_{jj}) \quad (4.3)$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \quad (4.4)$$

and an arithmetic mean for the repulsive exponent n_{ij}

$$n_{ij} = \frac{1}{2} (n_{ii} + n_{jj}) \quad (4.5)$$

The force field optimization was carried out by minimizing the objective function

$$f(\mathbf{p}) = \frac{1}{N^{\text{exp}}} \sum_{n=1}^{N^{\text{exp}}} \left(\frac{\Omega_n^{\text{sim}}(\mathbf{p}) - \Omega_n^{\text{exp}}}{\Omega_n^{\text{exp}}} \right)^2 \quad (4.6)$$

where Ω_n^{sim} and Ω_n^{exp} are simulated and quasi-experimental observables, respectively, containing liquid density and vapor pressure data, $\Omega \in \{\rho_L, p^{\text{sat}}\}$ of several substances. Quasi-experimental data were generated with DIPPR correlations in steps of 10 K in the temperature range between approximately $0.58 \cdot T_c$ and $0.95 \cdot T_c$. Vector \mathbf{p} contains the adjustable parameters, $\mathbf{p} = (\varepsilon_{ii}, \sigma_{ii}, n_{ii}, q_i)$, where index i is in this case a generic index that can, for example, represent a CH_3 -group as well as a CH_2 -group, when alkanes are considered. In previous works the Mie parameters (σ_{ii} and ε_{ii}) were optimized simultaneously supported by the PC-SAFT model[1, 3, 6, 107]. Further the parameters n_{ii} were optimized according to the same objective function, although discretized to even integer values. Similarly, for polar substances, the point charges were discretized along a predefined grid and the optimization of Mie parameters was done for every charge.

Intramolecular interactions like bond lengths, angle potentials and torsional potentials were taken from existing force fields, specifically the TraPPE force field [50, 53, 54] and the OPLS force field [44]. Fixed bond lengths between interaction sites are assumed.

4.2.2. The iTAMie force field

The concept of individualizing a transferable force field is introduced for ensuring good correlation results for well-known substances. The description of these substances with transferable force fields is sub-optimal, because the molecular model is adjusted as the

4. Individualized force fields based on the transferable anisotropic Mie potential

Table 4.1.: Pair potential parameters of UA-groups: energy parameter ε , size parameter σ , repulsive exponent n , and point charge q for the TAMie force field

pseudoatom i	$M_i/(\text{g/mol})$	$\varepsilon_i/k_B/\text{K}$	TAMie			ref.
			$\sigma_i/\text{\AA}$	n_i	charge q_i/e	
CH ₃ (sp ³)	15.035	136.318	3.6034	14		[1]
CH ₂ (sp ³)	14.027	52.9133	4.0400	14		[1]
CH (sp ³)	13.015	14.5392	4.3656	14		[157]
CH ₂ (sp ²) olefins	14.027	100.681	3.6005	14		[1]
CH (sp ²) olefins	13.018	53.9515	3.8234	14		[1]
O aldehydes	15.999	100.85	3.0276	12	-0.422	[3]
CH _x neighbor aldehydes					-0.038	[3]
CH (sp ²) aldehydes	13.019	68.934	3.4941	12	0.46	[3]
O ketones	15.999	65.55	3.093	12	-0.49	[3]
C (sp ²) ketones	12.011	32.775	3.919	12	0.49	[3]
O ether	15.999	86.7	2.84	12	-0.30	[6]

Table 4.2.: Bond lengths used in the proposed TAMie force field taken from literature.

type	$r_0/\text{\AA}$	ref
CH ₃ -CH _x $x \in \{1, 2\}$	1.54 + 0.2	[1, 6]
CH _x -CH _y $x, y \in \{1, 2\}$	1.54	[50]
CH ₃ -CH[=O _{ald}]	1.52 + 0.2	[3]
CH ₂ -CH[=O _{ald}]	1.52	[53]
CH ₃ -C[=O _{ket}]	1.52 + 0.2	[3]
CH ₂ -C[=O _{ket}]	1.52	[53]

Table 4.3.: Bending angles and constants for the proposed TAMie force field taken from literature.

bending site	$\theta/^\circ$	$k_0/k_B/\text{K}/\text{rad}^2$	ref
CH _x -CH ₂ -CH _y	114.0	62 500	[50, 136]
CH _x -C _{ket} -CH _y	117.2	62 500	[53]
C _{ket} -CH ₂ -CH _x	114.0	62 500	[53]
CH _x -C _{ket} =O _{ket}	121.4	62 500	[53]
CH _x -CH=O _{ald}	121.4	62 500	[53]

Table 4.4.: Torsional potential constants taken from literature.

torsion sites	$c_0/k_B/\text{K}$	$c_1/k_B/\text{K}$	$c_2/k_B/\text{K}$	$c_3/k_B/\text{K}$	ref
CH _x -CH ₂ -CH ₂ -CH _y	0.0	355.03	-68.19	791.32	[44, 50]
CH _x -CH ₂ -C=O _{ket}	2035.58	-736.90	57.84	-293.23	[53]
CH _x -C-CH ₂ -CH _y	-17.26	752.60	14.89	282.10	[54]
C-CH ₂ -CH ₂ -CH _x	0.0	355.03	-68.19	791.32	[44, 50]
CH _x -CH ₂ -CH ₂ -CH[=O _{ald}]	11.81	467.80	-274.10	846.80	[53]
C-CH ₂ -CH=O _{ald}	1182.0	-225.60	302.80	-339.30	[54]

4. Individualized force fields based on the transferable anisotropic Mie potential

best compromise in correlating force field parameters to many substances simultaneously. For the TAMie force field the most pronounced deviations are observed for vapor pressures (of polar substances), as visualized in Fig. 4.1.

We introduce the individualization parameter ψ_A that scales all van der Waals energy parameters ε_{ii} of a substance A according to

$$\varepsilon_{ii}^{\text{ind}} = \psi_A \cdot \varepsilon_{ii} \quad \forall ii \quad (4.7)$$

Index ii runs over all ((the mathematical symbol \forall says for *for all*)) interaction sites of substance A . For determining ψ_A -values we follow the same optimization procedure as described for the TAMie force field parameter. The objective function $f(\mathbf{p})$ now only depends on scalar-valued ψ_A as degree of freedom (instead of a parameter-vector \mathbf{p} of higher dimensionality) because all species A are decoupled and can be considered individually. The starting value for ψ_A is unity, leading to results of the original TAMie parameterization. We use the PC-SAFT model to approximate and minimize the objective function $f(\mathbf{p})$ as described by van Westen et al.[107] and Hemmen and Gross[1], leading to a next iteration value of ψ_A . As a convergence criterion, we considered the minimization satisfactory when for two subsequent iteration steps, $n - 1$ to n the decline of the objective function is below 0.05% ($f^{(n)}(\psi_A) - f^{(n-1)}(\psi_A) \leq 0.05\%$). Convergence is usually observed after one to two iterations.

4.2.3. Simulation Details

All simulations were carried out using a monte carlo algorithm in grand canonical ensemble with transition matrix sampling[76, 78, 140], and multiple histogram reweighting[13, 72, 73, 75] in a post processing step. The simulation method is in detail described in two previous papers[1, 6]. The chosen volumes vary between 50,000 \AA^3 for smaller molecules like propane or diethyl-ether to 160,000 \AA^3 for tetracosane. For all substances the number of molecules in the simulation volume reached around 300. Only for hexadecane to tetracosane the maximum numbers were around 200 molecules.

The sampling of insertion, deletion, or regrowth steps was enhanced using the configurational-biasing (CB) method [83, 158], with one CB-trial steps for each \langle interaction site \rangle at the lowest densities up to eight CB-trial steps per interaction site at highest densities. Further simulation details are summarized in the supporting information.

4. Individualized force fields based on the transferable anisotropic Mie potential

Table 4.5.: Comparison of different methods for individualizing the TAMie force field for two substances (butanone and di-n-pentyl-ether).

scaling method	butanone AAD/%	di-n-pentyl-ether AAD/%	average AAD/%
vdW energy (ψ_A -value)	1.25 (1.012)	2.1 (0.991)	1.7
vdW size (ψ_A^σ -value)	2.5 (1.010)	2.1 (0.9985)	2.3
partial charges (ψ_A^q -value)	1.26 (1.025)	5.0 (0.0)	3.1

4.3. Results

4.3.1. Assessment of the scaling method

This study investigates a simple (one-parameter) concept of individualizing transferable force fields for substances that are comprehensively characterized by experimental data.

First, we show that scaling the Mie energy parameter according to Eq. (4.7) is a suitable concept. We compare the approach with two alternative scaling schemes, namely scaling the Mie size parameter according to $\sigma_{ii}^{\text{ind}} = \psi_A^\sigma \cdot \sigma_{ii}$ and scaling the electrostatic potential as $q_i^{\text{ind}} = \psi_A^q \cdot q_i$. The comparison is made for two representative substances (butanone and di-n-pentyl-ether). In contrast to our objective function which is defined as the sum of squared deviations (RMSD), we chose absolute average deviation values (AAD values) for all comparisons of our results to quasi-experimental data, since AADs are much more intuitive to interpret and assess. Table 4.5 lists absolute average deviations (AAD) for all three approaches and shows that scaling the van der Waals energy parameter is indeed most effective for reducing the deviations in the properties captured in the objective function. It is interesting to note that the deviations for di-n-pentyl-ether (dpe) can not be reduced below 5.0 % using the electrostatic scaling, because the value of the scaling reaches the lower bound of zero ($\psi_{dpe}^q = 0$). A graphical representation of the sensitivities of how the objective function changes with the scaling parameters is given in the supporting information. With the results of Table 4.5 we conclude that scaling the Mie energy parameter is the most promising approach if a transferable force field should be individualized with a single parameter.

We individualized the TAMie force field with correction parameters ψ_A according to Eq. (4.7) for various substances of different chemical families: alkanes, olefins, ethers and ketones. The correction parameters and the corresponding absolute average deviations (AAD) are summarized in Table F.4. The table also lists deviations for the original TAMie force field (corresponding to $\psi_A = 1$). The considered objective function for optimizing the parameters ψ_A includes liquid densities and vapor pressures of the substances for reduced temperatures $T_R = T/T_C$ between $T_R \approx 0.57 - 0.96$. Fig. F.4 and F.7 show that individualizing the force field by adjusting ψ_A parameters leads to a markedly improved description of vapor pressures both in AD and in AAD values. The AD and AAD values

4. Individualized force fields based on the transferable anisotropic Mie potential

capture deviations between quasi-experimental data (DIPPR[137, 138]) and results from MC simulations. The improvement of the individualized TAMie to original TAMie is mild for (linear) apolar species and is particularly pronounced for polar substances. The liquid densities are only slightly improved (Table F.4). Values of AD (Fig. F.7) are useful measures for identifying systematic deviations, where the considered properties (vapor pressure and liquid density) are uniformly over- or underpredicted. For assessing the ability of a model to approximate experimental data, AAD values, however, are more meaningful and we restrict consideration to AAD-values henceforth. Tabulated AD values and graphical representations of AD values are shown in the supporting information. Graphical representations of Table F.4 are available in the supporting information.

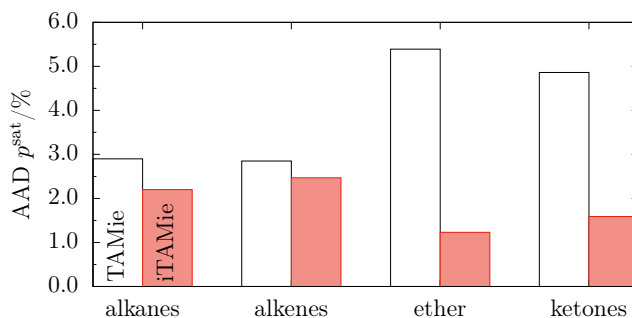


Figure 4.2.: Individualization of the TAMie force field decreases the absolute average deviation for calculated vapor pressures from quasi-experimental data[137] significantly.

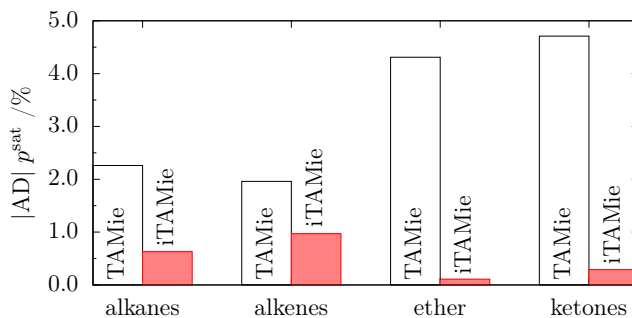


Figure 4.3.: Individualization of the TAMie force field also decreases the absolute value of the average deviation for calculated vapor pressures from quasi-experimental data[137] significantly.

4.3.2. Alkanes and olefins

The correction parameters ψ_A for alkanes and olefins differ only slightly from unity. The most considerable corrections of the Mie energy parameter are about 0.4 % (corresponding to $\psi_A = 1.0 \pm 0.004$) for branched alkanes and about 0.8 % for 5-methyl-1-hexene. The TAMie force field already shows good agreement with experimental data for many alkanes and olefins[1], so that the individualization for these substances does not lead to a significant improvement. Substances in Table F.4 are marked bold if we consider the individualization useful. For most alkanes and olefins the original TAMie force field parameters can be used without the individualization parameter ψ_A .

4.3.3. Ethers

The smallest member of a homologous series in TAMie is usually parameterized individually. That is the case for dimethyl-ether[6]. Other ethers are described assuming transferability of the force field parameters. We individualized diethyl-ether, di-n-propyl-ether and ethyl-propyl-ether as well as di-n-butyl and di-n-pentyl-ether. Figure 4.4a and 4.4b show the relative deviations from quasi-experimental data[137] in vapor pressure and liquid density, respectively. The diagrams show results for diethyl-ether and dipropyl-ether obtained for the TAMie model and for the individualized TAMie force field. The empty symbols were shown in Fig. 4.1. One observes significant improvement in the correlation of vapor pressure (Fig. 4.4a), whereas deviations in liquid densities remain almost unaltered (Fig. 4.4b). In comparison to well-established force fields one might consider the errors of the original TAMie model for liquid density to be relatively high. These deviations are a result of the chosen objective function which captures vapor pressures and liquid densities with equal weights. The final (dis)agreement to the experimental data is entirely defined by the objective function. The concept of individualizing with a single parameter does not alleviate these deviations significantly. Of course, if we had used a lower weight for vapor pressure within the objective function in the outset, the resulting force field would show a much improved agreement to liquid densities.

In the process of optimizing ethyl-propyl-ether we found that the correlation of the DIPPR data base for vapor pressure lacks accuracy for higher temperatures. The DIPPR correlation for vapor pressure is proposed for the temperature range 145.65 K to 500.23 K, although only values from 246.54 K to 359.96 K were used to fit the parameters of the DIPPR correlation. We decided to refit the vapor pressure curve considering all available experimental data points[159–163]. The result is shown in Fig. 4.5. The absolute average deviation between experimental data and the correlations thereby reduces from

4. Individualized force fields based on the transferable anisotropic Mie potential

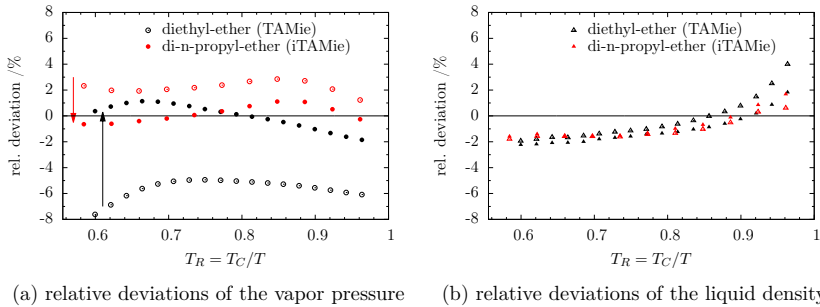


Figure 4.4.: Deviations in vapor pressure of the iTAMie force field (filled symbols) and of the TAMie force field (open symbols) from quasi-experimental data, for diethyl-ether (black symbols) and di-n-pentyl-ether (orange symbols).

$AAD_{DIPPR} = 2.81\%$ to $AAD_{new} = 1.48\%$ with

$$AAD = \frac{1}{N} \sum_1^N \left| \frac{(p_{exp}^{sat} - p_{cor}^{sat})}{p_{cor}^{sat}} \right|. \quad (4.8)$$

We repeated the optimization of the ψ_A -parameter for individualizing the TAMie force field, represented by the filled spheres in Fig. 4.5. The improvement is visible especially for higher temperatures. The fitting equation and the resulting parameters are given in the supporting information.

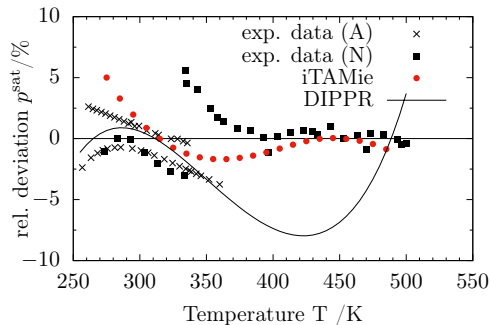


Figure 4.5.: Relative deviation for the vapor pressure of ethyl-propyl-ether in comparison to our newly adjusted correlation function. For this empirical correlation function all experimental data points [159–163] (crosses and filled squares) were used. Crosses represent the data points used for the original DIPPR correlation (A) and filled squares (N) were additionally used by us. For comparison, the relative deviation of the DIPPR correlation to the newly adjusted correlation is also given.

4.3.4. Ketones

We investigated the behavior of five different ketones. The absolute average deviation (combined p^{sat} and ρ^{liq}) reduces from 3.4 % to 1.75 % when applying the correction parameters ψ_A to the ketones. Figure 4.6 illustrates the relative deviations in vapor pressure (spheres) and liquid density (triangles) for butanone and 2-pentanone before and after the individualization. The figure shows the increased accuracy in vapor pressure, while the liquid densities have only little variation. In numbers, the errors due to vapor pressure reduce from 9.27 % to 0.32 % for butanone and from 2.68 % to 1.68 % for 2-pentanone. Because we chose a single parameter ψ_A to reduce systematic offsets of calculated vapor pressure, the slope of the vapor pressure curve is not much affected. In order to keep the approach simple, we made no attempt to further improve the representation of vapor pressure or liquid density by introducing more degrees of freedom. The relative deviations of the liquid density for butanone mildly decrease from 2.56 % to 2.17 %, whereas a slight increase is seen for 2-pentanone, from 2.25 % to 2.46 %. For example, the vapor pressure for butanone is overestimated when using the TAMie force field ($\psi_A > 1$), while it is underestimated for longer ketones ($\psi_A < 1$).

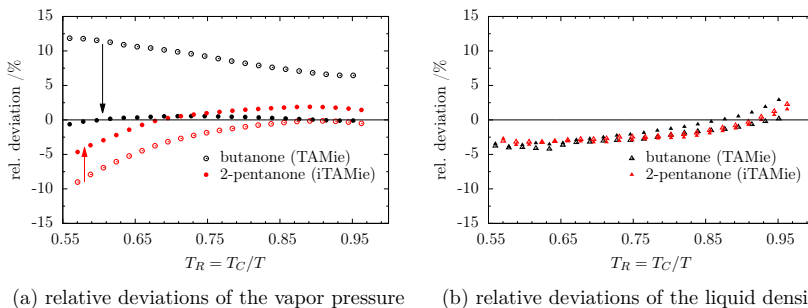


Figure 4.6.: Improved description of the vapor pressure for ketones when using the individualized force field. Filled symbols represent simulation results of the iTAMie force field while open symbols represent the TAMie force field. Triangles and spheres denote relative errors of the liquid density and vapor pressure, respectively.

4.3.5. Application to a binary mixture

To clarify the benefit of individualizing the TAMie force field, we apply iTAMie to a binary mixture. Figure 4.7 shows the phase diagram of butanone and 1-hexene at $T = 333.15K$. The calculation relies on the simple combining rules, Eq. (5.3) - (5.5), so that no binary cross-potential parameters were adjusted. Both substances have individualized parameters, which leads to different values of the vapor pressure for the

4. Individualized force fields based on the transferable anisotropic Mie potential

Table 4.6.: Optimization parameters ψ_A and AAD of TAMie and iTAMie compared to DIPPR correlations for different substances. Ethyl-propyl-ether* has been adjusted to quasi-experimental data of a newly fit correlation function. We consider individualizing useful for substances marked in bold font, whereas for most alkanes and olefins the improvement lies within the combined uncertainty of experimental data and simulations.

substance	AAD TAMie /%	AAD iTAMie /%	parameter ψ_A
propane	0.79	0.51	0.9988
butane	1.42	0.56	1.0024
pentane	0.87	0.80	1.0007
hexane	0.62	0.55	1.0004
heptane	1.02	0.96	1.0003
octane	1.28	1.20	1.0010
nonane	1.36	1.29	1.0009
decane	1.64	1.55	1.0016
undecane	2.12	2.04	1.0000
dodecane	2.5	2.5	1.0000
hexadecane	3.51	3.22	1.0022
eicosane	4.57	3.77	1.0027
tetracosane	3.79	3.14	1.0023
2-methylpentane	1.93	1.11	0.9968
3-methylpentane	2.22	0.60	1.0039
2-methylhexane	1.93	1.62	0.9977
2-methylheptane	1.77	1.69	0.9990
1-pentene	1.25	0.85	0.9978
1-hexene	1.86	1.67	0.9986
1-octene	2.08	1.68	0.9976
1-nonene	2.27	1.62	0.9960
1-decene	2.75	2.00	0.9955
1-undecene	2.55	2.49	0.9969
1-dodecene	2.64	2.38	0.9974
1-tridecene	2.89	2.86	0.9985
4-methyl-1-pentene	2.44	2.44	1.0000
5-methyl-1-hexene	4.33	1.92	0.9926
diethyl-ether	3.95	1.16	0.9931
ethyl-propyl-ether	2.22	2.22	1.0000
ethyl-propyl-ether*	2.09	1.29	0.9972
di-n-propyl-ether	1.93	0.94	1.0024
di-n-butyl-ether	2.28	1.23	0.9969
di-n-pentyl-ether	5.69	2.11	0.9908
butanone	5.91	1.25	1.0119
2-pentanone	2.46	2.07	0.99671
2-hexanone	2.27	1.57	0.9967
2-heptanone	2.97	2.19	0.9949
2-octanone	3.60	1.67	0.9939

4. Individualized force fields based on the transferable anisotropic Mie potential

pure substances in comparison to TAMie. Because the correction parameter for 1-hexene is close to unity, the right boundary of Fig. 4.7 does not change considerably, whereas for the left boundary we find a significantly improved representation of the vapor pressure of pure butanone. For correlating mixtures with high accuracy (and for improved subsequent predictions), it is usually required to adjust cross-van der Waals parameters, beyond the simple combining rules applied here. The example of Fig. 4.7 clearly reveals that meaningful binary interaction parameters for phase equilibria can only be determined if the pure component vapor pressures are reproduced in sufficiently good agreement to experimental data.

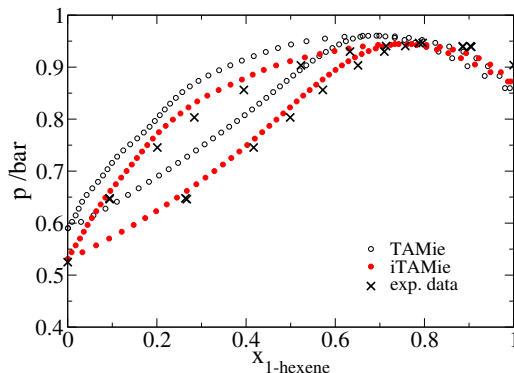


Figure 4.7.: Binary mixture of butanone and 1-hexene at a temperature of $T = 333$ K. Open spheres are results determined with TAMie, filled spheres are results from the iTAMie force field, and crosses are for experimental data of Hanson and van Winkle^[164].

4.4. Conclusion

In this work we showed that the individualization of a transferable force field is a promising approach. A single scaling parameter is introduced, acting on the van der Waals energy. The approach significantly reduces errors in vapor pressure for pure substances. This is especially true for polar substances such as ethers and ketones. For alkanes and olefins, where the original TAMie model is already in good agreement to experimental data, the improvement is marginal and the individualized results may lie within the combined uncertainties of simulations and experimental data. The so-obtained results make predictions and correlations of molecular simulations particularly interesting for applications in chemical engineering, where the vapor pressure is a critical quantity. The better agreement for pure component allows adjusting more meaningful binary interaction parameters for mixtures and should lead to better predictions of mixture phase equilibria.

5. Phase equilibria of binary mixtures with alkanes, ketones, and esters based on the transferable anisotropic Mie force field

The content of this chapter is a literal quote of the publication

D. Weidler and J. Gross. Phase equilibria of binary mixtures with alkanes, ketones, and esters based on the transferable anisotropic mie force field. Fluid Phase Equilib., 490:123–132, 2019

J. Gross had the role of a daily supervisor and was involved in editing the manuscript. Additions or deletions compared to the published work are marked with angular brackets.

The transferable anisotropic Mie potential (TAMie) is extended to acetates and formates. The parameters defining the Mie pair potential, which represent the van der Waals interactions as well as point charges were optimized, following an approach of Hemmen and Gross [J. Phys. Chem. B, 2015, 119, 11695-11707]. Intramolecular bending and torsion potentials were obtained from quantum chemical calculations. Binary mixtures of ethyl acetate/hexane and, because mixture data of esters is scarce, butanone/n-hexane, butanone/n-heptane, 2-pentanone/n-heptane and were examined with the TAMie and the iTAMie force field. The cross interaction parameters for mixtures of alkanes and ketones were determined to $\kappa_{CH_2O=} = 0.0306$ and $\kappa_{CH_3O=} = 0.0542$. The mixture of ethyl acetate and heptane shows an azeotropic point and is predicted very well from simulations in grand canonical ensemble using the iTAMie force field as comparison with experimental data shows. Additionally to phase equilibria, we calculated excess properties for the mixture of hexane and butanone, which also show very promising results compared to experimental data.

5.1. Introduction

The accurate extrapolation or prediction of phase equilibria for mixtures is of central importance for the design and optimization of chemical processes and process equipment. Molecular simulations with transferable force fields are a powerful tool for predicting phase equilibria and other thermodynamic properties of pure substances and of mixtures. Several force fields have been developed with emphasis on thermodynamic properties and phase equilibria. Among the most established models are the Transferable Potentials for Phase Equilibria (TraPPE) [40, 50–52, 90, 112, 165], the Optimized Potential for Liquid Simulations (OPLS) [44–49] and the Assisted Model Building with Energy Refinement (AMBER) force field [19]. The parameters of OPLS were chosen to represent experimental properties of liquids. Such properties are e.g. the enthalpy of vaporization and the liquid density. AMBER stands for a family of force fields. The main purpose is to enable the modelling of biomolecules, where most investigations focus on conditions around 298 - 313 K and 1 bar, often in aqueous solution. The TraPPE force field was developed to describe and predict phase equilibria requiring at least a temperature range from the triple point to the critical point. Emphasis was placed on the representation of coexisting liquid densities ρ^L for given temperature T leading to excellent representations of pure T - ρ projections or of T - x projections, where x_i denotes the mole fraction of species i in a mixture, but the force field shows somewhat higher deviations of vapor pressure.

The aforementioned force fields are based on Lennard-Jones pair potentials for modelling van-der-Waals interactions and fixed point charges for approximating the electrostatic structure of the molecules. Additional degrees of freedom can be introduced to the parametrization of a force field by using an ansatz function for intermolecular potentials other than the Lennard-Jones potential, e.g. the Mie potential with three adjustable parameters per van der Waals interaction site [111, 133, 149, 150] or the Buckingham potential also with three adjustable parameters [114]. Toxvaerd [117, 147] proposed to move the interaction site of, say, a CH_3 group away from the position of the carbon atom to effectively account for the presence of hydrogen atoms. The distance of moving united-atom groups, outwards from the position of the large atom introduces additional degrees of freedom in the parameterization of a force field. This concept led Ferrando, Boutin, Ungerer and coworkers to the Anisotropic United Atom (AUA) force field [123–131, 148].

The Transferable Anisotropic Mie (TAMie) force field describes molecules as Mie interaction sites (with defined attractive exponent of $n = 6$) and with fixed point charges. The development of TAMie is based on a well-defined objective function, where the parameterization was conducted simultaneously to several members of a chemical family [1, 3, 6]. The objective function captures squared deviations in calculated vapor pressure and liquid densities to experimental data. Meaningful predictions were achieved for phase equilibria

ria of mixtures. Inherent to a force field, where parameters defining interaction sites are assumed to be transferable from one substance to another substance (and where the objective function includes several substances of a chemical family), the force field parameterization is a compromise leading to more pronounced deviations for each substance, as compared to force fields individually optimized for a considered substance [166]. To alleviate this problem while maintaining the advantage of transferable force field, namely the ability to predict pure component and mixture behavior of systems weakly characterized by experimental data, we proposed an individualized version of the TAMie force field [4], referred to as iTAMie.

In this study we extend the TAMie force field by proposing parameters for esters. The force field is subsequently applied to binary mixtures containing esters. Because experimental data of mixtures with esters are scarce, we also investigate mixtures containing ketones. Predicted phase equilibria of mixtures are in rather satisfying agreement to experimental data using Berthelot-Lorentz combining rules [109, 110]. To further improve the results we determined transferable cross-energy parameters for mixtures of esters with alkanes and of ketones with alkanes.

5.2. Force field development

In this section we first describe the development of the intramolecular parameterization of the TAMie force field. The bond-angle bending and torsional potentials of ester groups are determined from *ab initio* calculations of ethyl formate and ethyl acetate. Subsequently, the development of van der Waals and coulombic parameters is presented.

5.2.1. *Ab initio* calculations of intramolecular bending and torsional potentials

In previous work on the TAMie force field [1, 3, 6] bending potentials and torsional potentials as well as most of the bond-to-bond distances between different united atom groups were taken from existing, well established force fields, mostly from the TraPPE [40, 50–52, 90, 143, 165, 167] and the OPLS force field [44–48]. For esters a variety of intramolecular potentials is available in literature. Kamath et al. extended the TraPPE force field to esters [167]. For methyl acetate these authors conducted quantum chemical scans of the potential energy surface at the HF/6-31g+(d,p) level for the $\text{CH}_3\text{-C-O-CH}_3$ and O=C-O-CH_3 torsional potential. The adjusted potentials smoothly follow these energy surfaces. The torsional potentials, however, are not independent from each other. When applying this parameterization all energies are twice as high as expected, because two torsion potentials contribute to the energy, whereas the parameterization (we suspect)

projected the total energy on each one torsional potential.

Maerzke et al. recently published torsional potentials for the two dihedrals [168]. They used electron structure density functional theory with the B3PW91 functional and with the 6-31++G(d,p) basis set. Because they split the total energy symmetrically to both dihedrals, their potential function is not applicable to formates using a united-atom force field. Our goal is to adjust force field parameters for esters, including formates, which is why we performed MP2 structure optimization using a TZVPP basis set and constraints for ethyl formate and ethyl acetate. The calculations were conducted using the quantum chemistry program package TURBOMOLE [169] for the structure optimizations. Figure 5.1 shows the relative energy differences between the optimal structure of ethyl formate and different predefined dihedral angles of the O=CH-O-CH₂ torsion obtained with the MP2 structure optimization. The resulting energy function, with torsional parameters listed in Table 5.3, is almost identical to the energy calculated with the TraPPE potential function, except that our approach is only based on one dihedral in a united-atom model. We adjusted a torsional potential function to the energies observed from the MP2 calculations of ethyl formate and apply these parameters to all other formates.

Figure 5.2 gives the energy surface of the CH-O-CH₂-CH₃ dihedral of ethyl formate and the C-O-CH₂-CH₃ dihedral of ethyl acetate in comparison to the torsional potential used by TraPPE [165], which has its origin in OPLS [49]. Because our quantum chemical calculations showed a different behavior of this dihedral, we adjusted a new potential function to ethyl formate. Quantum chemical scans for ethyl acetate confirm that the potential function can also be applied to other esters.

Figure 5.3 shows the energy landscape for different dihedral angles of the O=C-O-CH_x torsion of acetates. The value of 0 ° corresponds to the *cis* configuration. The O=C-O-CH_x dihedral and the CH_y-C-O-CH_x dihedral share the same axis (C-O) and the central carbon atom is sp² hybridized, so that we assume a 180 ° phase shift between these two dihedrals and adjust our potential function with equal weights to both dihedrals. For comparison the potential functions of TraPPE for esters [167] and for acrylates [168] are also shown. The parameters of all potential functions are listed in Table 5.3.

5.2.2. Optimization of intermolecular force field parameters for esters

In agreement to earlier work on the TAMie force field [1, 3, 4, 6], the pair potential between two interaction sites located on different molecules, or positioned more than three bonds apart within the same molecule, is given as the a Mie potential plus an electrostatic

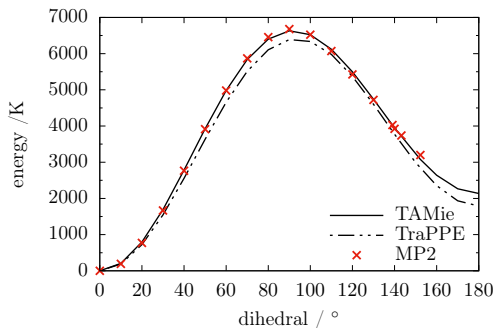


Figure 5.1.: Torsional energies for the $\text{O}=\text{CH}-\text{O}-\text{CH}_2$ dihedral of ethyl formate calculated with the MP2 approach (crosses) and the proposed torsional potential (solid line) of TAMie. For comparison the torsional energies from TraPPE [168] (dashed dotted line) are included. However, in TraPPE these energies are a result of two combined dihedrals. Because there is only one dihedral along the $\text{CH}-\text{O}$ axis in a united atom model of formates, we adjusted a new potential function.

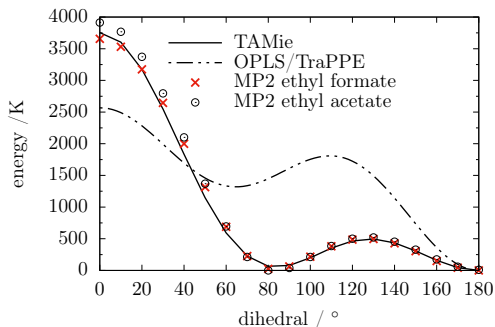


Figure 5.2.: Torsional energies for the $\text{CH}_x-\text{O}-\text{CH}_2-\text{CH}_3$ dihedral angle in esters calculated with the MP2 approach (red crosses) for ethyl formate (index $x = 1$) and the proposed torsional potential (black solid line). The MP2 calculation for ethyl acetate (index $x = 0$) is also shown (black spheres). For comparison the potential function of the TraPPE force field [165] is given. This torsional potential has its origin in the OPLS model [49] and was adjusted to ethyl methyl ether.

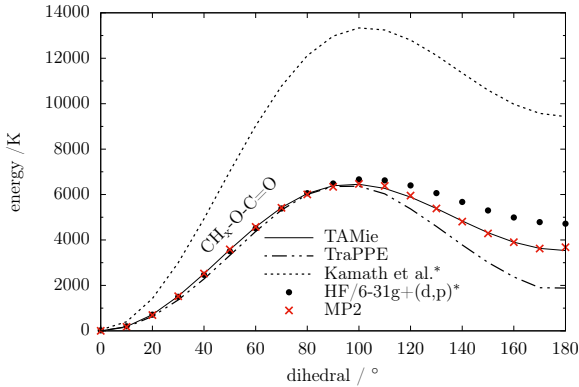


Figure 5.3.: Torsional potentials for ethyl acetate. Comparison of *ab initio* calculations on MP2 level (red crosses) to the adjusted potential functions of TAMie (black solid line). The dotted line is the potential function of Kamath et al. [167] and the dashed dotted line is the TraPPE potential function [168]. The energies of all potential functions are the result of two combined dihedrals, namely of $\text{O}=\text{C}-\text{O}-\text{CH}_2$ and of $\text{CH}_3-\text{C}-\text{O}-\text{CH}_2$ with a phase shift of 180° between them according to the sp^2 hybridization of the central carbon atom. The black filled circles show the energy obtained by Kamath et al. [167] through quantum chemical optimization of methyl acetate.

potential, as

$$u_{ij} = c(n_{ij}) \cdot \varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{n_{ij}} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi \varepsilon_0 r_{ij}} \quad (5.1)$$

with r_{ij} as the distance between two interaction sites i and j , ε_{ij} is the van der Waals energy parameter, σ_{ij} denotes the corresponding size parameter. Further, n_{ij} is the repulsive exponent, and q_i is the partial point charge. The partial charge or the energy parameter ε_{ij} can be zero for a given pair of interaction sites i and j . Further, ε_0 is the vacuum permittivity and the constant $c(n_{ij})$ is defined to ensure the minimum of the Mie potential at $-\varepsilon_{ij}$, with

$$c(n_{ij}) = \left(\frac{n_{ij}}{n_{ij} - 6} \right) \left(\frac{n_{ij}}{6} \right)^{\frac{6}{n_{ij}-6}} \quad (5.2)$$

For cross-wise parameters of the Mie potential we adopt the Lorentz-Berthelot combining rules [109, 110]

$$\sigma_{ij} = \frac{1}{2} (\sigma_{ii} + \sigma_{jj}) \quad (5.3)$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii} \varepsilon_{jj}} \quad (5.4)$$

5. Phase equilibria of binary mixtures with alkanes, ketones, and esters based on TAMie

and an arithmetic mean for the repulsive exponent n_{ij} , as

$$n_{ij} = \frac{1}{2}(n_{ii} + n_{jj}) \quad (5.5)$$

Like in previous work on the TAMie force field [1, 3, 6] we treated the effective two-body van-der Waals and the effective (non-polarizable) point charges as adjustable parameters. Parameters of esters are perceived as transferable with both ester-oxygen and with the connecting ester-carbon atom as van der Waals interaction sites. Because the smallest members of a homologous series often have non-transferable force field parameterizations, we consider the subgroup of formates (i.e. all methanoates) through an individual set of TAMie parameters, whereas all other esters are treated as transferable.

To obtain an optimal set of parameters we used the method described by van Westen et al. [107] and modified by Hemmen and Gross [1]. This procedure uses the Perturbed-Chain Statistical Associating Fluid Theory (PC-SAFT) equation of state to approximate the objective function of the force field optimization. Applying PC-SAFT is useful because the multidimensional parameter optimization within the analytic model is conducted in milliseconds and is thus much faster than performing a full force field optimization based only on (many subsequent) molecular simulations. The force field parameters so determined from PC-SAFT are used as a next iteration step with molecular simulations. The optimization procedure using an analytic equation of state (PC-SAFT) is iterative because the equation of state only approximates the true objective function when extrapolating force field parameters. Once convergence is achieved, however, the analytic equation of state does not act on the objective function, as shown by Hemmen and Gross [1], and the procedure thus minimizes the actual objective function. The optimization procedure was used to obtain van der Waals parameters of the esters. Point charges were changed along a predefined grid of values, with values around the point-charge values from the TraPPE force field. For each value of the point charge along the predefined grid, we optimized the size parameters σ_{ii} and energy parameters ε_{ii} . In order to reduce the number of degrees of freedom in this multi-dimensional optimization problem, we introduced fixed ratios for σ_{ii}/σ_{jj} and $\varepsilon_{ii}/\varepsilon_{jj}$. These ratios were taken from the TraPPE force field [167], defined as: $\sigma_{-(H)C}/\sigma_{-O} = 1.3643$, $\sigma_{=O}/\sigma_{-O} = 1.0893$, $\varepsilon_{-(H)C}/\varepsilon_{-O} = 0.7273$ and $\varepsilon_{=O}/\varepsilon_{-O} = 1.4364$. The index -(H)C implies that we took these ratios both for formates and acetates.

The optimization minimizes the objective function

$$f(\mathbf{p}) = \frac{1}{N^{\text{exp}}} \sum_{n=1}^{N^{\text{exp}}} \left(\frac{\Omega_n^{\text{sim}}(\mathbf{p}) - \Omega_n^{\text{exp}}}{\Omega_n^{\text{exp}}} \right)^2 \quad (5.6)$$

where Ω_n^{sim} and Ω_n^{exp} are simulated and experimental observables, respectively, and N^{exp} is the number of experimental data points. We use liquid density and vapor pressure data

as observables in the objective function, $\Omega \in \{\rho^L, p^{\text{sat}}\}$, with equal weight between them. In the objective function, we consider liquid density and vapor pressure data of several substances simultaneously. By regarding several species within the optimization problem, we ensure a balanced compromise in the resulting force field parameters.

Force field parameters of formates and of acetates were optimized independent of one another. Ethyl acetate, propyl acetate, butyl acetate and pentyl acetate were used for the optimization of acetates. And ethyl formate, propyl formate and butyl formate were defined as the training set for the optimizing the formate-parameters of the TAMie force field.

5.3. Simulation details

5.3.1. Pure substances

Simulations for pure substances were performed in the grand canonical ensemble, using histogram reweighting in a post-processing step. In the grand canonical ensemble the temperature T , volume V and chemical potential μ are fixed whereas the number of molecules N fluctuates. We use the transition matrix method [76, 78, 140] to determine a bias potential on the fly, ensuring an approximately equal sampling of N . Furthermore, we divide the N -space into several windows and run a separate simulation of each window. All simulations run in parallel and only require an overlap between two neighboring windows. If the first window samples N between 0 and 10, for example, then the second window samples between $N = 10$ and 20. The bias potential from transition-matrix sampling ensures that the overlapping molecule number between two windows ($N = 10$ in the example) is well-sampled from both sides, making wider overlaps and weighting schemes, such as the Ferrenberg-Swendsen approach [73] unnecessary. Any attempt of inserting a molecule beyond the upper bound of a window is trivially rejected (and analogously for the lower bound). Using the multiple ensemble technique [141] these windows can be combined to a total histogram. A detailed description of the simulation techniques can be found in previous works [1, 6].

The volume of the simulation box varied between 50,000 Å³ for ethyl formate and 80,000 Å³ for pentyl acetate, corresponding to about 300 molecules for the highest densities of each substance. Probabilities for the different moves were set to: particle insertion/deletion 40%, translational and rotational displacement 50% and particle regrowth 10%. The cut-off radius is $R_c = 14$ Å and analytic tail corrections are applied. For coulombic interactions we apply standard Ewald summation with the damping parameter $\alpha = 7.59/L$ and the maximum number of k -vectors set to $k_{\text{max}} = 9$ in each direction. Esters are 'branched' molecules and require careful generation of trial configura-

rations for growing a new molecule, for regrowing an existing molecule, and for determining the Rosenbluth weight when trial-deleting a molecule in a configurational biasing scheme [51, 83, 158]. The branching node with neighboring interaction sites is defined as a molecular fragment, similar to how Shah and Maginn define molecular fragments [87]. We apply a Monte Carlo scheme with importance sampling on isolated molecules to generate representative trial configurations [170]. The generation of these fragments occurs during runtime up to a maximum number of $N^{\text{fragment}} = 5 \cdot 10^5$ equilibrated, representative configurations. The coordinates of these branching nodes (fragments) are stored in a simple array in random-access memory. After N^{fragment} fragments are stored, we do not continue to generate new configurations, but randomly draw branching nodes from the existing $5 \cdot 10^5$ fragments, which significantly accelerates the calculation. The structures of a branching node are generated using the correct probability distribution of angle potentials and only the torsional potential (and intermolecular potentials) need to be assessed in a CBMC trial configuration. This approach is similar to methods proposed in refs [87, 89, 114]. The MC code is written and maintained by members of our research group. A comparison to results from Cassandra [171], an established open source code, was given in a previous study [3]. Further simulation details, such as the number of equilibration and production steps and the number of configurational biasing steps are given in the supporting information.

5.3.2. Binary mixtures

The simulation of binary mixtures mainly follows the scheme proposed by Shen and Errington [79] where simulations in grand canonical ensemble with transition matrix methods are carried out to calculate phase coexistence properties of binary Lennard-Jones mixtures.

Similar to the case of pure substances, we divide the simulation domain into different windows of a fixed size $\Delta N_1 \times \Delta N_2$. For every window a Monte Carlo simulation in grand canonical ensemble is executed with five different moves: particle insertion 20%, particle deletion 20%, translational displacement 25%, rotation 25% and particle regrowth 10%. The cut-off distance and parameters of the Ewald summation are the same as detailed for pure substances. A configurational bias scheme [51, 83, 158] in conjunction with first storing branching nodes in an array (library) and later randomly drawing from that library [87, 89, 114] improves the statistics of insertion, deletion and regrowth moves with increasing step numbers for higher densities. A histogram $P(N_1, N_2)$, i.e. a relative probability distribution of molecule numbers, is the result of every simulated window. To get the overall probability density $P(N_1, N_2)$ the histogram of all windows are stitched together according to a scheme described in ref. [6].

5.3.3. Cross-energy parameters κ_{ij} for binary mixtures

The cross-wise potential between two van der Waals united-atom sites can be approximated using Lorentz-Berthelot combining rules, according to Eq. (5.3) and (5.4). In fact, the Berthelot-Lorentz combining rules have gone into the parameterization of the pure substances. For mixtures, we allow the cross-energy parameter ε_{ij} between united-atom group i of one species and group j of another species to deviate from the Berthelot rule. We express the cross-energy parameter as the Berthelot value minus a correction, according to

$$\varepsilon_{ij} = (\varepsilon_{ii} + \varepsilon_{jj})^{0.5} (1 - \kappa_{ij}) \quad (5.7)$$

These cross-energy parameter κ_{ij} can be determined with help of the PC-SAFT equation of state. The connection between κ_{ij} and the binary interaction parameter $k_{\alpha\beta}$ of PC-SAFT is described by Schacht et al. [108] and by Waibel et al. [58], where α and β are different substances. (We note that in the original articles [58, 108] parameter $k_{\alpha\beta}$ was referred to as k_{ij} because in these references i and j indicate different substances). To achieve meaningful cross-energy parameters, one can consider several mixtures containing the same chemical families simultaneously in the objective function, and optimizing for κ_{ij} . For the optimization, we use PC-SAFT, as described in previous work [58, 108].

5.4. Results

This study proposes a transferable force field for esters, with a parameter-set for formates and another parameter-set for other esters. First, the intramolecular potential is described. Then, the optimization of force-field parameters of pure esters is detailed. Subsequently the phase behavior of mixtures of n -alkanes with esters and of alkanes with ketones is investigated. For an improved representation of substances with well-known properties, we furthermore individualize the TAMie force field and we adjust binary cross-energy parameters κ_{ij} . In addition to phase equilibrium properties of mixtures we calculate excess properties of mixtures in the liquid phase and compare results of the TAMie force field to experimental data.

5.4.1. Intramolecular potential

Table 5.1 lists the parameters and literature references for the bond lengths. The bond length between a $-\text{CH}_3$ -group and a neighboring carbon-group is in the TAMie force field extended by 0.2\AA compared to the carbon-carbon atom distance, to better account for the hydrogen in the $-\text{CH}_3$ -group. Tables 5.2 and 5.3 list the parameters and literature references for bending and torsional angles, respectively. Some values in these tables are taken from literature, while others were adjusted to *ab initio* calculations as described in

5. Phase equilibria of binary mixtures with alkanes, ketones, and esters based on TAMie

Table 5.1.: Bond lengths of the TAMie force field for esters.

Group 1	Group 2	Distance /Å	reference
-C(H)_ester	-O_ester	1.344	[167]
=O_ester	-C_ester	1.200	[167]
-CH ₃	-C_ester	1.54 + 0.2	[1]
-CH ₂	-C_ester	1.54	[1]
-CH ₂	-O_ester	1.41	[167]

Table 5.2.: Angle bending parameters of the TAMie force field for esters.

Type	$\theta_0 / ^\circ$	$k_0/k_b / \text{K/rad}^2$	reference
O-CH ₂ -CH _x (ester)	108.8	88418.0	this work
CH _x -CH ₂ -C (ester)	114.0	62500.0	[50, 136, 165]
C-O-CH _x (ester)	115.0	62500.0	[167]
O=C-CH _x (ester)	115.0	62500.0	[167]
O-C-CH _x (ester)	110.0	70600.0	
O=C-O (acetates)	125.0	62500.0	[167]
[O=]CH-O-CH _x (formates)	116.4	72570.0	this work
O=CH-O (formates)	126.6	136890.0	this work

previous sections. For the torsional potential we apply three types of expansions, namely

$$\frac{U_{\text{tors}}}{k_B} = c_0 + c_1 \cos(\phi) + c_2 \cos(2\phi) + c_3 \cos(3\phi) \quad (5.8)$$

$$\frac{U_{\text{tors}}}{k_B} = c_0 + c_1 (1 + \cos(\phi)) + c_2 (1 - \cos(2\phi)) + c_3 (1 + \cos(3\phi)) \quad (5.9)$$

$$\frac{U_{\text{tors}}}{k_B} = \sum_{n=0}^3 c_n (\cos(\varphi + \pi))^n \quad (5.10)$$

which, for the shifting angle $\pi = 0$, can be related to Eq. (5.8). The second-last column of Table 5.3 allocates the torsional parameters to the appropriate torsional model, Eq. (5.8) to (5.10).

Table 5.3.: Torsional potential parameters of the TAMie force field for esters.

torsion sites	$c_0/k_b / \text{K}$	$c_1/k_b / \text{K}$	$c_2/k_b / \text{K}$	$c_3/k_b / \text{K}$	Eq.	reference
O=C-O-CH _x	2039.19	-808.02	-1155.28	-75.90	(5.8)	this work
CH _x -C-O-CH _y	2039.19	+808.02	-1155.28	+75.90	(5.8)	this work
O=CH-O-CH _x	2205.23	-907.66	2758.96	-212.79	(5.9)	this work
(H)C-O-CH ₂ -CH _x	0.0	1293.83	-899.46	583.80	(5.9)	this work
O-CH ₂ -CH ₂ -CH _x	0.0	176.62	-53.34	769.93	(5.9)	[165]

5.4.2. Pure esters

Pure formates

We first regard the chemical family of formates. The optimized parameters are $\sigma_{-HC=}$, $\sigma_{=O}$, σ_{-O-} as well as $\varepsilon_{-HC=}$, $\varepsilon_{=O}$, ε_{-O-} and partial charges. Whereas all σ_{ii} and ε_{ii} were optimized simultaneously, we considered varying partial charges in a range $q_{HC=} = 0.36$ to 0.50 in a predefined grid of 8 points and optimized all other parameters for each of the 8 grid points. Charge $q_{HC=}$ refers to the carbon atom with a double bond towards the neighboring ester oxygen atom (=O). The three other charges are then defined as $q_{=O} = -q_{HC=}$, $q_{O-} = -0.625 \cdot q_{HC=}$, and $q_{CH_x} = 0.625 \cdot q_{HC=}$, for the double and single bonded ester oxygen and for the neighboring CH_x -group, respectively.

The objective function (Eq.(5.6)) includes vapor pressure data and liquid density data of ethyl formate, propyl formate and butyl formate in a temperature range of $T/T_i^c = 0.56$ to $T/T_i^c = 0.95$, where T_i^c is the experimental critical temperature of the considered species. For each set of point charges the optimization of all van der Waals parameters was carried out automatized, i.e. the iterative procedure was automated where one iteration step comprises MC-simulations, histogram reweighting for phase equilibrium properties, and parameter optimization using PC-SAFT. Figure 5.4 presents the result (i.e. the lowest deviation) for varying point charges. For formates the lowest overall deviation was found for the point charge of the carbon site (next to the double bonded oxygen) as $q_{HC=}/e = 0.48$. Partial charges are given dimensionless, as factors of the unit electron charge e .

The objective function is defined as the sum of squared deviations (RMS). For assessing how the objective function varies with a force field parameter, as done in Fig. 5.4, it is useful to regard RMS values. However, for presenting results on single properties (such as vapor pressure or liquid densities) we give deviations as absolute average deviation (AAD) or as relative deviations from experimental data, because these measures are more tangible and intuitive to interpret. Relative deviations are used for graphical comparison of calculated results to experimental data, with relative deviation defined as $\frac{\Omega_n^{sim} - \Omega_n^{exp}}{\Omega_n^{exp}}$, where Ω_n as the n -th data point of the property of interest. AAD-values and RMS-values are defined in the supporting information. Figure 5.5 shows the vapor pressures of three formates with $q = 0.48$ in a Clausius-Clapeyron diagram and compares them to quasi-experimental data [137]. Because errors are not easily assessable in this representation, Figure 5.6 shows the same results as relative deviations of the vapor pressure for the three formates. Ethyl formate and propyl formate are found in good agreement with quasi-experimental data [137], except at low temperatures where the absolute relative deviations increase. The vapor pressure of butyl formate, however, is underestimated over the entire temperature range. Figure 5.7 shows that liquid densities are underestimated

for lower temperatures and overestimated near the critical point. The combined AADs of vapor pressures and liquid densities are 1.3 % for ethyl formate, 2.0 % for propyl formate and 2.7 % for butyl formate. AADs for the liquid densities lay between 0.8 % and 1.1 %. For vapor pressures the AADs are 2.0 % for ethyl formate, 3.3 % for propyl formate and 4.2 % for butyl formate. Although the description of the vapor pressure of butyl formate has a significant error compared to DIPPR data, the combined overall deviation (Eq. (5.6)) is minimal with this set of adjustable parameters.

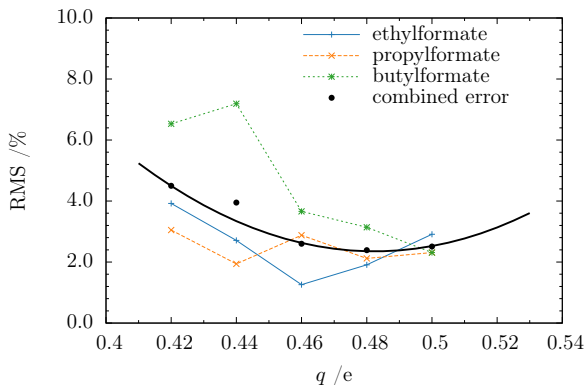


Figure 5.4.: Root mean square (RMS) deviations for formates of both, calculated vapor pressure p^{sat} and liquid density ρ^{L} from experimental values. Black spheres represent averaged RMS-values for all three formates, each for optimized van der Waals parameters. Black solid line is regression curve. The minimum for the combined optimization therefore is at $q_{\text{HC}=} = 0.48e$.

Pure acetates and other esters

The force field parameters assigned to all esters other than formates were adjusted to four substances: ethyl acetate, propyl acetate, butyl acetate, pentyl acetate. Similar to the optimization of formates we adjusted $\sigma_{>C=}$, $\varepsilon_{>C=}$, and the partial charge $q_{iC=}$. The ratios between $\sigma_{>C=}$, $\sigma_{=O}$, σ_{-O-} and the ratios between $\varepsilon_{>C=}$, $\varepsilon_{=O}$, ε_{-O-} were fixed in order to reduce the dimensionality of the force field optimization problem to three dimensions.

Figure 5.8 presents the root mean square error of the four pure acetates that are member of the objective function for different point charges between $q_{iC=} = 0.34$ and $q_{iC=} = 0.44$. Each point represents an optimal value with respect to the van der Waals parameters $\sigma_{>C=}$ and $\varepsilon_{>C=}$ (and the other related van der Waals parameters). For $q_{iC=} = 0.38$ the combined error of vapor pressure and liquid density is minimal and we use this value henceforth.

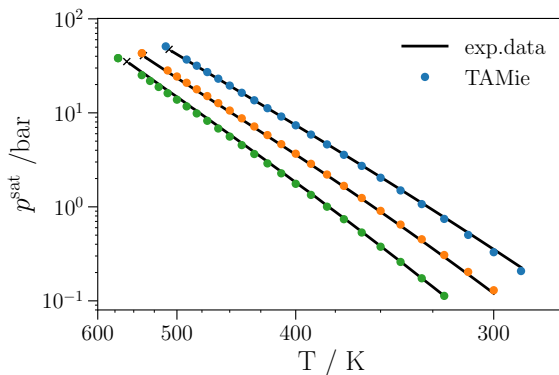


Figure 5.5.: Comparison of calculated vapor pressures from the TAMie force field to quasi-experimental data [137] for ethyl-formate (blue), *n*-propyl-formate (orange), *n*-butyl-formate (green).

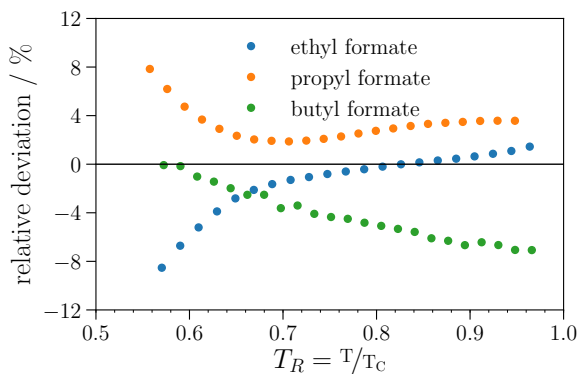


Figure 5.6.: Relative deviations of calculated vapor pressure from quasi-experimental data [137] for formates in the reduced temperature range from $T_R = 0.56 - 0.96$. Blue, orange and green spheres denote the relative deviations of ethyl-formate, *n*-propyl-formate and *n*-butyl-formate, respectively.

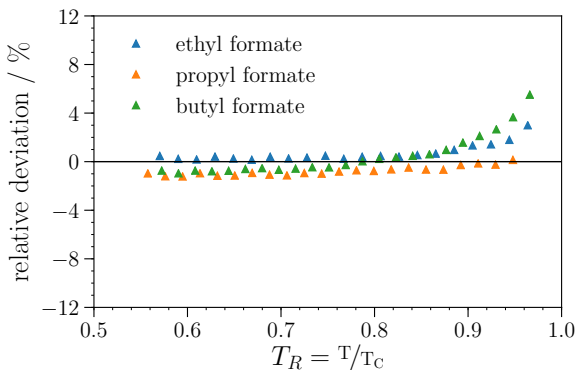


Figure 5.7.: Relative deviations of calculated liquid densities for formates from quasi-experimental data [137]. Blue, orange and green spheres denote the relative deviations of ethyl-formate, *n*-propyl-formate and *n*-butyl-formate, respectively.

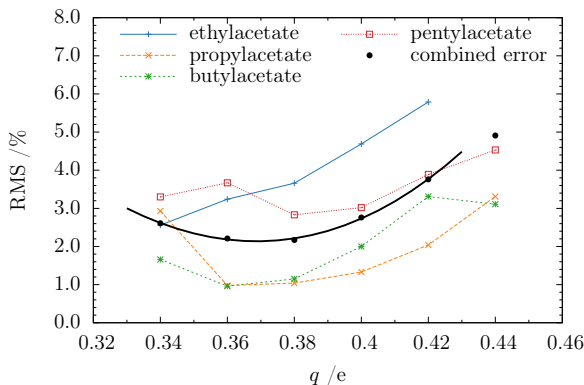


Figure 5.8.: Root mean square (RMS) deviations of both, calculated vapor pressure p^{sat} and liquid density ρ^L for acetates from experimental values. Black spheres represent averaged RMS-values for all four acetates, each for optimized van der Waals parameters. Black solid line is regression curve. The minimum for the combined optimization therefore is at $q_{iC=} = 0.38e$. Blue, orange, green and red lines are the root mean square errors of ethyl, *n*-propyl, *n*-butyl and *n*-pentyl acetate, respectively.

Table 5.4.: van der Waals (Mie) parameters and point charges of the TAMie force field for esters.

type	M_i / (g/mol)	σ / Å	ε / K	n_i	q / e	ref
-CH ₃ -[CH ₂]	15.035	3.6034	136.318	14		[1]
-CH ₂ -	14.027	4.0400	52.9133	14		[1]
-CH _x -[O_ester]				14	+0.2375	this work
-O_ester	15.999	2.727	33.66	12	-0.2375	this work
=O_ester	15.999	2.970	48.34	12	-0.380	this work
-C_ester	12.011	3.720	24.48	12	+0.380	this work
-CH _x -[O_formates]				14	+0.300	this work
-O_formates	15.999	2.619	68.461	12	-0.300	this work
=O_formates	15.999	2.853	98.334	12	-0.480	this work
-CH_formates	13.019	3.573	49.789	12	+0.480	this work

Figure 5.9 shows the relative deviation of the vapor pressure of 4 acetates compared to quasi-experimental data [137]. The simultaneous optimization of all four substances led to good agreement between calculated vapor pressures and experimental data for propyl and butyl acetate, whereas the vapor pressure for ethyl acetate is underestimated and for pentyl acetate it is underestimated for lower and overestimated for higher temperatures. The absolute average deviations (AADs) of the vapor pressures are 6.89 % for ethyl acetate, 0.91 % for propyl acetate, 1.13 % for butyl acetate and 2.43 % for pentyl acetate. The AADs for the liquid densities of the four acetates range between 0.38 and 0.71 %. Figure 5.10 gives a graphical analysis of the relative deviations of liquid densities confirming a rather good agreement to quasi-experimental data. (In this diagram, we use a refined scale for the relative deviations compared to Fig. 5.9, although both quantities enter with the same weight into the objective function). Figure 5.11 provides the T - ρ -projections for acetates.

5.4.3. Individualized pure esters

To overcome some shortcomings of transferable force fields, we recently introduced a correction parameter ψ_A acting on all van der Waals energy parameters ε_{ii} of one substance A. That leads to a more accurate description of the vapor pressures for well known substances while preserving the strength of a transferable force field that it can be applied to substances not well characterized by experimental data [4]. The individualization parameter ψ_A modulates the van der Waals energy parameters according to

$$\varepsilon_{ii}^{\text{ind}} = \psi_A \cdot \varepsilon_{ii} \quad \forall ii \quad (5.11)$$

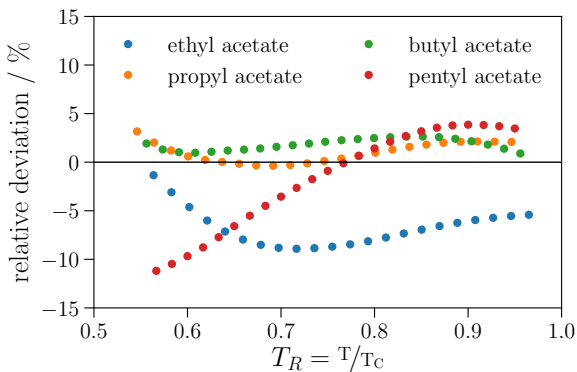


Figure 5.9.: Relative deviations of calculated vapor pressure from quasi-experimental data [137] for acetates in the reduced temperature range from $T_R = 0.57 - 0.96$. Blue, orange, green and red spheres denote the relative deviations of ethyl, *n*-propyl, *n*-butyl and *n*-pentyl acetate, respectively.

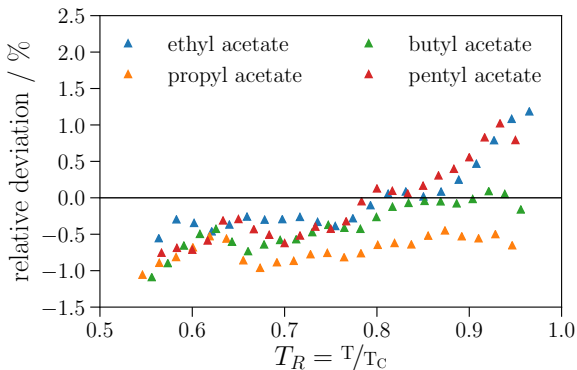


Figure 5.10.: Relative deviations of calculated liquid densities from quasi-experimental data [137] for acetates. Blue, orange, green and red spheres denote the relative deviations of ethyl, *n*-propyl, *n*-butyl and *n*-pentyl acetate, respectively.

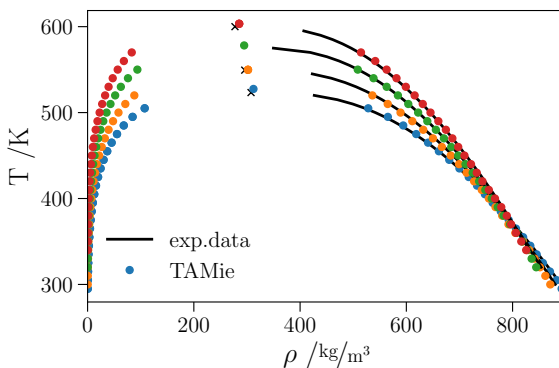


Figure 5.11.: Comparison of calculated phase equilibrium data from the TAMie force field to quasi-experimental data [137] in a T - ρ -projection. The black solid line are quasi-experimental data and blue, orange, green and red spheres denote the simulation results of n -pentyl-acetate, n -butyl-acetate, n -propyl-acetate and ethyl-acetate, respectively.

The index ii runs through all united atom groups of substance A and we briefly refer to this force field parameterization as individualized transferable anisotropic Mie (iTAMie) force field.

Following this scheme, we determined ψ_A for ethyl acetate to pentyl acetate and for ethyl formate to butyl formate. Table 5.5 gives the AAD values without ($\psi_A = 1$) and with individualization of the TAMie force field. Vapor pressure is most sensitive to parameter ψ_A and is most significantly improved through the individualization step. For acetates the improvement of p^{sat} is pronounced only for ethyl acetate, where the deviation in vapor pressure is reduced from 6.9 % to 3.3 % whereas the liquid densities are almost unaffected (0.4 to 0.6 %). For formates the improvement of p^{sat} is pronounced for butyl formate, where the deviation is reduced from 4.3 % to 1.0 % and it is moderate for propyl formate, with deviations in vapor pressure decreasing from 3.4 % to 2.0 %. The liquid densities for ethyl formate are slightly improved (0.7 to 0.3 %) whereas the liquid densities for propyl formate and butyl formate are unaffected within the statistical uncertainties.

5.4.4. Binary mixtures

We investigated the phase behavior of a mixture of ethyl acetate with heptane. Because experimental data for phase equilibria of mixtures with esters is scarce, we also study the phase behavior of three different alkane/ketone mixtures, namely butanone with n -hexane, butanone with n -heptane, and 2-pentanone with n -heptane.

Figure 5.12 shows a mixture of ethyl acetate and heptane at $T = 343.15$ K determined

5. Phase equilibria of binary mixtures with alkanes, ketones, and esters based on TAMie

Table 5.5.: Comparison of the combined absolute average deviations for the TAMie and iTAMie force field.

substance	AAD TAMie /%	AAD iTAMie /%	ψ_A
ethyl acetate	3.6	1.9	0.9923
propyl acetate	0.9	0.8	1.0012
butyl acetate	1.1	0.6	1.0021
pentyl acetate	2.4	2.4	1.0000
ethyl formate	1.4	1.3	0.9967
propyl formate	2.1	1.3	0.9862
butyl formate	2.7	1.9	0.9969

with the iTAMie force field. Using iTAMie (rather than TAMie) is advantageous because errors in the vapor pressure have a strong influence on the resulting phase diagram on the pressure-scale of this binary diagram. Even with the individualization the vapor pressure for ethyl acetate is slightly underestimated for this temperature, as seen on the left hand side of Figure 5.12. Overall, however, we consider the resulting phase diagram to be in rather good agreement with the experimental data, especially considering the fact that no cross-energy parameters κ_{ij} were adjusted to obtain the azeotropic behavior.

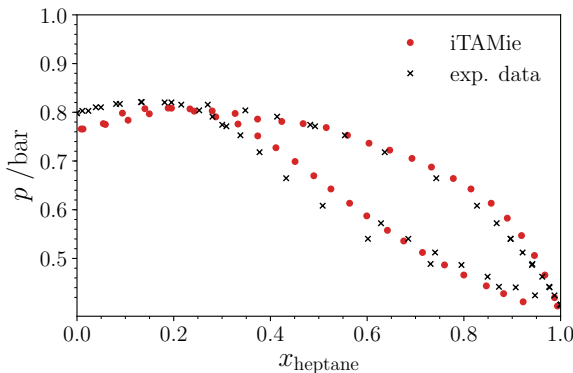


Figure 5.12.: Phase diagram of ethyl acetate with heptane at $T = 343.15$ K. Red filled circles represent results of the iTAMie force field (without corrections to the Berthelot-Lorentz combining rules) and black crosses show experimental data [172].

Experimental data for mixtures of esters with alkanes is scarce, which is why we also investigate mixtures of ketones with n -alkanes. The TAMie force field for ketones was earlier reported [3]. The phase diagram of butanone in mixture with n -hexane is shown in Fig. 5.13 for three levels of modelling: the topmost diagram shows predictions of the TAMie force field, with a satisfying agreement to experimental data. The vapor pressure

of pure butanone is somewhat underestimated. The individualization inherent to the iTAMie force field leads to an improved description of the vapor pressure of pure butanone, as the center diagram of Fig. 5.13 shows. The bubble-point pressure and dew point pressure is slightly underestimated. As a third level of modelling, we adjust binary cross-energy parameters by rescaling energy-parameters ϵ_{ij} as determined from the Berthelot combining rule with the factor $(1 - \kappa_{ij})$, according to Eq. (5.7). The lowest diagram of Fig. 5.13 shows the phase diagram of butanone with hexane with the resulting binary cross interaction parameters $\kappa_{CH_X,O} = \kappa_{CH_X,C(Ketone)} = 0.082$ where CH_X represents both, a CH_3 and a CH_2 -group. The results from the molecular simulations are thereby in very good agreement to the the experimental data. For adjusting meaningful cross interaction parameters κ_{ij} , it is of importance to represent vapor pressures of pure substances with sufficient accuracy.

In Figure 5.14 and 5.15 we investigate the transferability of force field parameters to mixtures of butanone with *n*-heptane and of 2-pentanone with *n*-heptane, respectively. Values of κ initially determined for the mixture of butanone with *n*-hexane were transferred to these two mixtures without further adjustment. Transferring binary parameters to butanone / *n*-heptane at somewhat lower temperatures is a rather small step and gives good agreement to experimental data, as Figure 5.14 shows. For the mixture of 2-pentanone with *n*-heptane one observes somewhat higher deviations of iTAMie to experimental data (Figure 5.15).

Figure 5.16 shows the excess enthalpy of the mixture butanone and hexane at $T = 298.15$ K calculated from the iTAMie force field in comparison to experimental data of Kiyohara et al. [175] and Murakami et al. [176]. This diagram suggests that good predictions of excess properties can be obtained for force fields parameterized to phase equilibrium properties. Good agreement for the excess enthalpy specifically implies that the temperature behavior of the phase equilibrium is robust for extrapolations in temperature, as the Gibbs-Helmholtz equation demands. Monte Carlo simulations in grand canonical ensemble with transition matrix and histogram reweighting is a suitable method to obtain excess quantities with rather low statistical uncertainty.

5.5. Conclusion

In this work we developed parameters of the TAMie force field for esters. A distinction is made between formates and other esters (whereby mainly acetates are considered). This study provides intramolecular potentials for esters determined from *ab initio* calculations. Van der Waals parameters and partial charges for the ester group were obtained by minimizing a clearly defined objective function defined for several substances simultaneously. Phase equilibria calculated with the proposed TAMie force field are in good agreement to

5. Phase equilibria of binary mixtures with alkanes, ketones, and esters based on TAMie

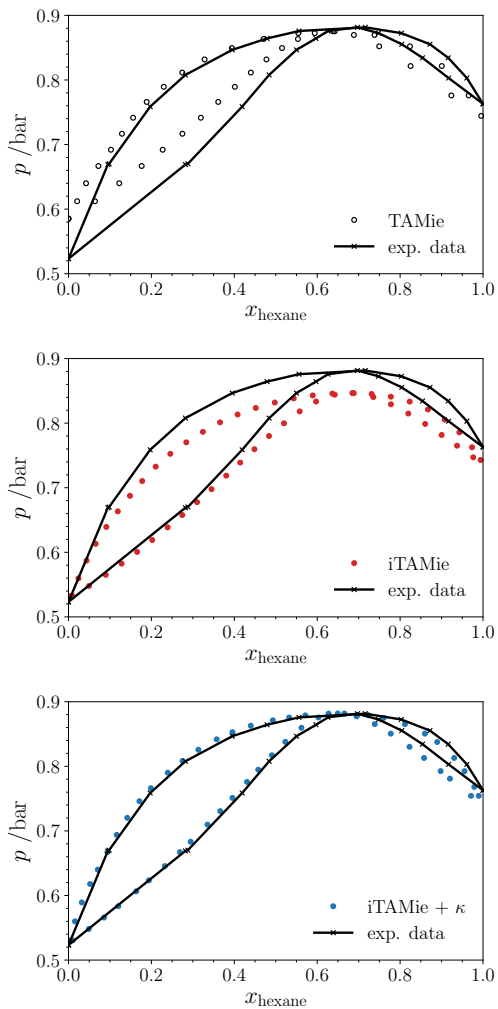


Figure 5.13.: Vapor liquid equilibrium of butanone with *n*-hexane at $T = 333.15$ K. Black symbols are results from the TAMie force field. Red spheres are results from the iTAMie force field. The blue filled spheres represent results from the iTAMie force field with adjusted cross-energy parameters $\kappa_{CH_X,O} = 0.082$ and $\kappa_{CH_X,C(Ketone)} = 0.082$ with $X \in \{2, 3\}$. Black crosses represent experimental data from [Hanson and van Winkle \[164\]](#).

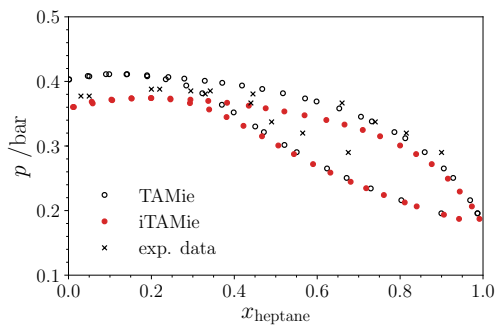


Figure 5.14.: Vapor liquid equilibrium of butanone with *n*-heptane at $T = 323.15$ K. Black symbols are results from the TAMie force field. Red spheres are results from the iTAMie force field using cross-energy parameters obtained from Fig. 13. Black crosses are experimental data from [Aristovich et al. \[173\]](#)

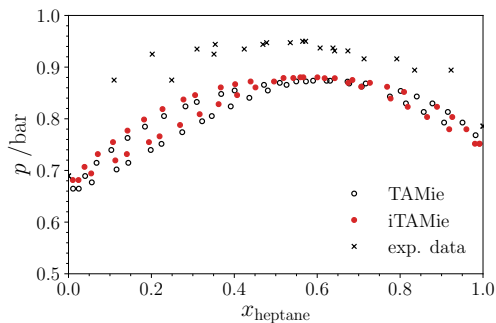


Figure 5.15.: Vapor liquid equilibrium of 2-pentanone with *n*-heptane at $T = 363.15$ K. Black symbols are results from the TAMie force field. Red spheres are results from the iTAMie force field using cross-energy parameters obtained from Fig. 13. Black crosses are experimental data from [Scheller and Rao \[174\]](#)

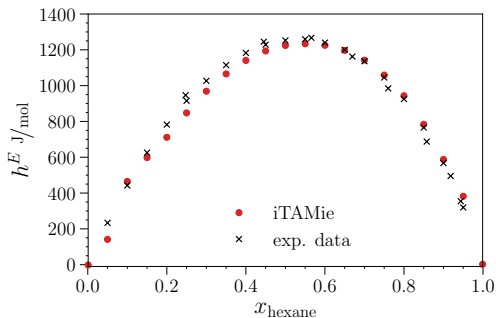


Figure 5.16.: Excess enthalpies h^E for the mixture of butanone and hexane at $T = 298.15$ K. There is good agreement between the experimental data [175, 176] (black crosses) and results from the iTAMie force field (red circles).

experimental data.

Phase equilibria and excess properties of mixtures were investigated. Good agreement of results from the TAMie force field to experimental data was observed for the mixture of ethyl acetate with hexane. For mixtures of ketones with n -alkanes the TAMie force field was also seen in good agreement to experimental data, especially once transferable cross interaction parameters were adjusted. The calculated excess enthalpy h^E of the system butanone/hexane closely follows experimental data, implying (through the Gibbs-Helmholtz relation) robustness in extrapolations to other temperatures.

6. Conclusion and outlook

This work presents new TAMie force field parameters for pure polar substances like ketones, aldehydes and esters as well as TAMie parameters for small cyclic alkanes. TAMie is a transferable united-atom force field. Force field parameters for intermolecular interactions (van der Waals parameters σ_{ii} and ε_{ii} and point charges) were optimized by minimizing an objective function containing experimental vapor pressures and liquid densities often using several members of a chemical family. Very good correlation results are obtained for these quantities (vapor pressure and liquid densities). Intramolecular force field parameters were either adjusted to quantum mechanical calculations (esters) or were taken from literature after assessing them through quantum mechanical calculations.

To speed up the minimization of the objective function, the PC-SAFT equation of state was used as a surrogate model as proposed by [Hemmen and Gross\[1\]](#). The procedure starts with conducting Monte Carlo molecular simulations. The PC-SAFT equation is explicitly expressed in terms of the force field parameters and the objective function is then minimized using the PC-SAFT model. The computational cost of this minimization step is low. To obtain the true minimum of the objective function, the sequence of conducting Monte Carlo simulations and minimizing the objective function using the PC-SAFT model is iterative. Usually, the optimum of the objective function was found after 3-5 iterations. PC-SAFT was also used to generate the input conditions for the molecular simulations, i.e. defining suitable (T, μ) -conditions to approximately trace the phase envelope of a substance.

The whole procedure of generating input conditions, examining molecular simulations in Grand Canonical ensembles, post-processing, minimizing the objective function and generating new input parameters, was automated by an MPI (Message Passing Interface) implementation (see Appendix A). For the Grand Canonical Monte Carlo simulations, the phase space was divided into several windows. Each window is considered as a single molecular simulation which needs a certain amount of CPU time to finish. For the post processing - i.e. combining and reweighting the histograms and calculating the substance properties - all simulations have to be combined. The duration of a single simulation, however, depends on the number of molecules in the window. For large molecule numbers (representing the liquid phase) the required simulation time is longer than for small molecule numbers (representing the vapor phase). The reserved cores were

6. Conclusion and outlook

utilized about 20 % more efficiently when two windows of different densities were run on the same core in sequence, compared to the trivial solution of running each window on a separate core.

Individual parameters were assigned and optimized for cyclopentane and cyclohexane. Absolute average relative deviations (AAD) in vapor pressure and liquid density are 0.5 % compared to experimental data. For ketones, as presented in chapter 3, minimizing the objective function for three ketones (butanone, 2-pentanone and 2-octanone) led to an AAD of 4.0 % (combined deviation of vapor pressures and liquid densities) and are dominated by the deviations in vapor pressure. Transferring these parameters to 3-pentanone and 2-hexanone yields a combined deviation (AAD) of 2.4 %, which confirms the transferability of the determined parameters. Taking all five simulated ketones into account, the average combined deviation (AAD) in vapor pressure and in liquid density is 3.5 %. The error in liquid density is high (2.2 %), compared to other force fields, however, the deviation in vapor pressure – compared e.g. to established transferable force fields like TraPPE – is significantly lower.

This work also introduces the concept of individualized transferable force fields (iTAMie), as presented in chapter 4. The concept offers the possibility to maintain the predictive character of a transferable force field for substances with scarce experimental data, while avoiding higher deviations of the force field with respect to experimental data for well-measured substances. Especially for polar substances such as ethers and ketones, the introduction of iTAMie reduced average deviations in vapor pressures from 5.4 % to 1.3 % and from 4.8 % to 1.6 %, respectively. This feature is also useful when predicting the behavior of mixtures. It was shown, that this approach works for determining phase equilibria of mixtures, by applying it on the (azeotropic) mixture of butanone and 1-hexene. With the individualized force field, the mixture properties are in good agreement with experimental data.

In chapter 5 transferable force field parameters for esters (split into the subgroups of formates and other esters) were developed with combined AADs of 2.1 % for formates and 2.0 % for acetates. Mixtures of ketones with alkanes were investigated and it was shown, that the combination of an individualized transferable force field and transferable cross-energy parameters $\kappa_{\alpha\beta}$ is a promising concept. A cross-energy parameter was determined for the system of *n*-hexane and butanone and applied to mixtures of butanone with *n*-heptane as well as 2-pentanone with *n*-heptane. The results of all three mixtures show good agreement with experimental data.

After completing the study, it was found that for substances with multiple interaction sites per molecule the choice of the energy bin-width has an influence on simulation results at low temperatures. The impact of the energy bin-width is larger than can be explained by statistical uncertainty and should be further clarified in subsequent work (Appendix

6. Conclusion and outlook

D).

A. CPU time usage

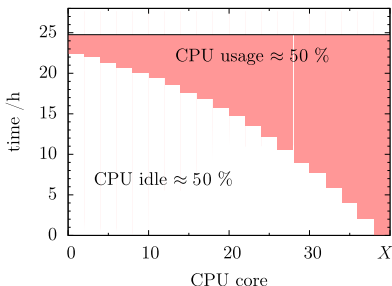
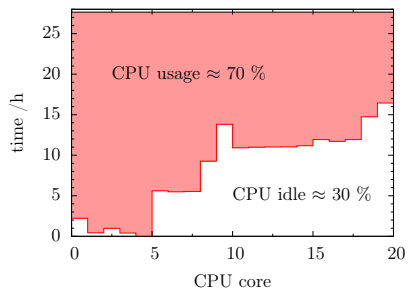


Figure A.1.: Expected CPU time usage when using MPI to handle the different windows in multi-canonical Monte Carlo when each window runs on a single CPU

The purpose of the MPI implementation is to automatize the optimization of force field parameters. Since Monte Carlo simulations still require a lot of computational resources, simulations were carried out on a compute cluster. To use this cluster efficiently, all used CPU cores should have similar workloads. In all GCMC simulations with multiple histograms in this work the the N-space is divided into different windows of size ΔN , where the first window starts from $N = 0$ (ideal gas) and the last window represents a dense fluid. Details were given in section 2.2.1. The volume of all simulations is equal. If each CPU core did the calculation of a single window, each core would have a different runtime, resulting in an inefficient use of computational resources (Fig. A.1) The simplest way to improve that behaviour, is to successively run two windows on every core. The window with the least number of particles ran on the same CPU as the window with the highest number of particles. If the windows are indexed from 0 to X , the first CPU executes 0 and X , the second 1 and $X - 1$, the third 2 and $X - 2$ and so on. The first index is zero in agreement with the standard in MPI implementations. Figure A.2 shows the resulting usage, which shows an improvement of about 20% compared to the simplest distribution scheme.

A. CPU time usage

Figure A.2.: CPU time usage when using MPI to handle the different windows in multicanonical Monte Carlo and two windows run on each CPU.



B. TAMie for benzene

Simulations to optimize the force field parameters for benzene have been done in grand canonical ensemble with transition matrix and histogram reweighting, analogously to all previous described optimizations. Benzene, as the simplest molecule of aromatic hydrocarbons, is build up as a ring of six carbon atoms a hydrogen atom bound to each carbon atom. According to the united atom approach the carbon atom together with the hydrogen atom are treated as one interaction site. I assume a rigid, planar structure and the angles between three neighboring united-atom sites are 120° . Figure B.1 shows the structure of the molecule. Because of this structure, only the parameters $\sigma_{ij}, \varepsilon_{ij}$ for

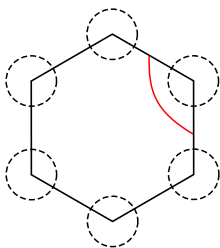


Figure B.1.: Structure of a united-atom benzene molecule. Schematic depiction of the planar ring with united-atom interaction sites -CH and the stiff bending angle θ (in red).

the united-atom group -CH and the repulsive exponent $n_{ij} = n$ of the Mie potential (see equation (B.1)) were optimized, for the Mie potential.

$$u_{ij} = c(n_{ij}) \cdot \varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{n_{ij}} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (\text{B.1})$$

Figures B.2 and B.3 show the relative deviations of vapor pressure and liquid density of the optimized TAMie parameters for different exponents of the Mie potential compared to DIPPR correlation functions[137]. It is obvious, that for a Mie exponent of $n = 16$ the overall deviation is minimal in this case, since both, vapor pressure and liquid densities show the lowest deviations compared to quasi experimental data. In numbers, the absolute average deviation of the liquid density is 1.0 % and the AAD of the vapor pressure is 0.3 %. Table B.1 shows the optimized TAMie parameters.

B. TAMie for benzene

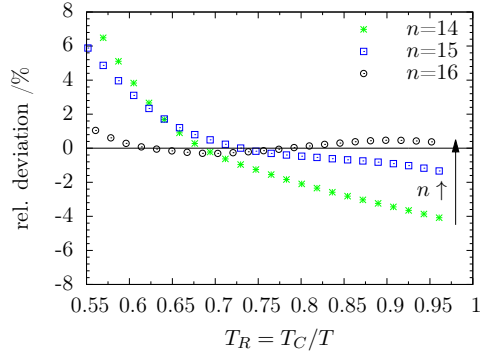


Figure B.2.: Relative deviations of the vapor pressure of benzene with different repulsive exponents n compared to quasi-experimental data of DIPPR[137].

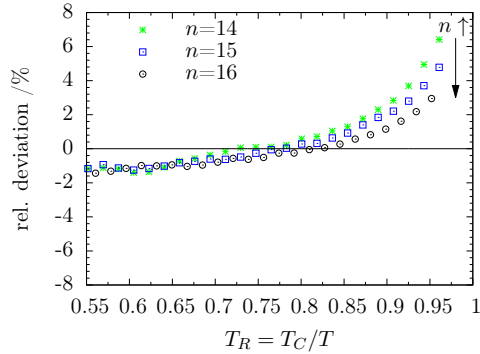


Figure B.3.: Relative deviations of the liquid density of benzene with different repulsive exponents n . The deviations are calculated against DIPPR correlation functions.

Table B.1.: van der Waals (Mie) parameters, bond length l and bending angle θ of the TAMie force field for benzene.

type i	$l / \text{\AA}$	$M_i / (\text{g/mol})$	$\sigma_{ii} / \text{\AA}$	$\varepsilon_{ii} / \text{K}$	n_i	$\theta / ^\circ$
$-\text{CH}_{\text{benz}}$	1.40	13.019	3.7163	66.238	16	120

B.1. Dynamic properties of benzene

The dynamic properties in this chapter were determined by the group of J. Vrabec at the University of Paderborn in 2018. They developed their own model of benzene [177], a Lennard-Jones force field with quadrupoles (LJ-Q) and did the simulations with their codes and the parameters of this work. The comparison between the two models in Figures B.4-B.6 show very good agreement, both between the two models and between the models and the experiments or DIADEM correlations of the DIPPR database [137] fit to experimental data, respectively.

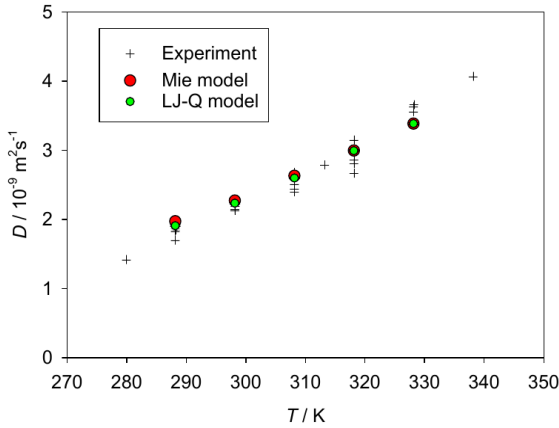


Figure B.4.: Diffusion coefficients of benzene calculated with the TAMie force field and the LJ-Q force field [177]. All results for molecular simulations (MD) were obtained by J. Vrabec. Experimental data from DDB[139, 178–181]

The correlation used for quasi-experimental data of the shear viscosity of benzene is

$$Y = \exp\left(A + \frac{B}{T} + C \cdot \ln(T) + D^E\right) \quad (\text{B.2})$$

with $A = 7.5117$, $B = 294.68$, $C = -2.794$ and $D = E = 0$ and T inserted in Kelvin[137].

The correlation used for quasi-experimental data of the thermal conductivity of benzene is

$$Y = \frac{A}{B^{(1+(1-T/c)^D)}} \quad (\text{B.3})$$

with $A = 0.0542518$, $B = 2.74187$, $C = -7.22561$ and $D = 8.22561$ and T inserted in Kelvin[137].

B. TAMie for benzene

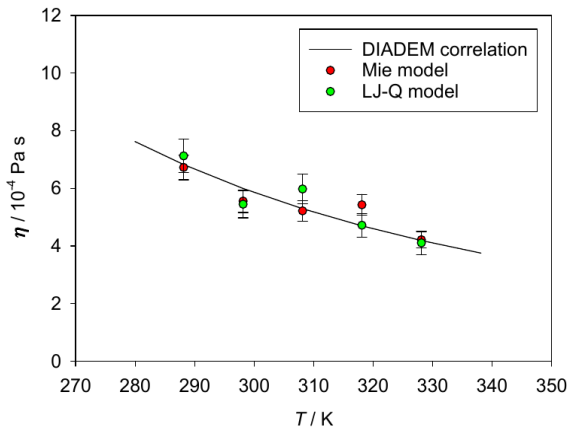


Figure B.5.: Shear viscosity of benzene calculated with the TAMie force field and the LJ-Q force field [177]. All results for molecular simulations (MD) were obtained by J. Vrabec. Quasi-experimental data from DIADEM correlations of the DIPPR database [137]

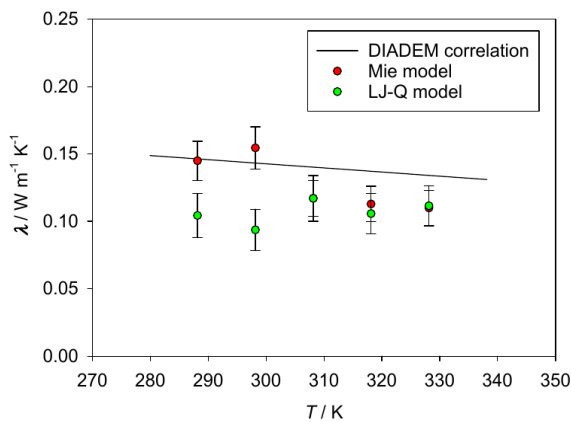


Figure B.6.: Thermal conductivity of benzene calculated with the TAMie force field and the LJ-Q force field [177]. All results for molecular simulations (MD) were obtained by J. Vrabec. Quasi-experimental data from DIADEM correlations of the DIPPR database [137]

B. TAMie for benzene

These results show, that it is possible to achieve very accurate dynamic properties by adjusting force field parameters only to liquid density and vapor pressure data. The TAMie force field for benzene works without coulombic interaction, i.e. no point-charges, dipolar or quadrupolar interactions.

C. Expressing the PC-SAFT model in terms of force field pure parameters

For the optimization of TAMie force field parameters the PC-SAFT equation of state is used. Therefore the following relations between force field parameters to PC-SAFT parameters (\hat{m} , $\hat{\sigma}$, $\hat{\varepsilon}$) are introduced and assumptions are made as proposed by [van Westen et al.\[107\]](#). What is needed to relate force field parameters to PC-SAFT parameters are group-contribution segment numbers m_α as taken e.g. from [Sauer et al.\[182\]](#).

The segment parameter \hat{m} :

$$\hat{m} = \sum_{\alpha}^{\text{seg.typ.}} N_{\alpha} m_{\alpha} \quad (\text{C.1})$$

As a first approximation it is assumed that the "volume" of a molecule within PC-SAFT is equal to the "volume" of the same molecule of the TAMie force field.

$$\frac{\pi}{6} \hat{m} \hat{\sigma}^3 \approx \frac{\pi}{6} \sum_{\alpha}^{\text{seg.typ.}} N_{\alpha} m_{\alpha} \sigma_{\alpha\alpha}^3 \quad (\text{C.2})$$

Because this does not apply exactly, the correction parameter ϕ_{σ} is introduced.

$$\hat{m} (\phi_{\sigma} \hat{\sigma})^3 = \sum_{\alpha}^{\text{seg.typ.}} N_{\alpha} m_{\alpha} \sigma_{\alpha\alpha}^3 \quad (\text{C.3})$$

The van der Waals parameter $\varepsilon_{\alpha\beta}$ appears in the dispersion term of the Helmholtz energy, which to first order reads

$$F^{disp} = 2\pi\rho N \sum_{\alpha}^{\text{seg.typ.}} \sum_{\beta}^{\text{seg.typ.}} N_{\alpha} N_{\beta} \int g_{\alpha\beta}(r) u_{\alpha\beta}^{disp}(r) r^2 dr, \quad (\text{C.4})$$

This perturbation approach can be applied to the PC-SAFT equation of state (left) as well as to the molecular model of the force field (right), leading to the first-order equality.

$$\sum_{\hat{\alpha}=1}^{\hat{m}} \sum_{\hat{\beta}=1}^{\hat{m}} \int \hat{g}_{\hat{\alpha}\hat{\beta}}^{hc}(r) \hat{u}^{disp}(r) r^2 dr = \sum_{\alpha}^{\text{seg.typ.}} \sum_{\beta}^{\text{seg.typ.}} N_{\alpha} N_{\beta} \int g_{\alpha\beta}^{fs}(r) u_{\alpha\beta}^{disp}(r) r^2 dr \quad (\text{C.5})$$

C. Expressing the PC-SAFT model in terms of force field pure parameters

By introducing a mean radial pair distribution function for hard chains

$$\bar{g}^{hc} = \frac{1}{\hat{m}^2} \sum_{\alpha=1}^{\hat{m}} \sum_{\beta=1}^{\hat{m}} \int \hat{g}_{\alpha\beta}^{hc}(r) \quad (\text{C.6})$$

one obtains

$$\hat{m}^2 \int \bar{g}^{hc} \tilde{u}^{disp}(r) r^2 dr = \sum_{\alpha}^{\text{seg.typ.}} \sum_{\beta}^{\text{seg.typ.}} N_{\alpha} N_{\beta} \int g_{\alpha\beta}^{fs}(r) u_{\alpha\beta}^{disp}(r) r^2 dr \quad (\text{C.7})$$

The equation above can be made dimensionless by introducing a reduced radial distance and a reduced radial distance \tilde{r} and a reduced potential function \tilde{u}^{disp} , if the force field considers a Lennard-Jones or Mie potential for the dispersive interactions. For PC-SAFT (left-hand side) these are defined as $\tilde{r} = r/\hat{\sigma}$ and $\tilde{u}^{disp}(\tilde{r}) = \tilde{u}^{disp}(r)/\hat{\varepsilon}$ and for TAMie (right-hand side) $\tilde{r} = r/\sigma_{\alpha\beta}$ and $\tilde{u}^{disp}(\tilde{r}) = u_{\alpha\beta}^{disp}(\tilde{r})/\varepsilon_{\alpha\beta}$. This results in an equation, where the Mie and PC-SAFT parameters are outside the integrals

$$\hat{m}^2 \hat{\sigma}^3 \hat{\varepsilon} \int \bar{g}^{hc}(\tilde{r}) \tilde{u}^{disp}(\tilde{r}) \tilde{r}^2 d\tilde{r} = \sum_{\alpha}^{\text{seg.typ.}} \sum_{\beta}^{\text{seg.typ.}} N_{\alpha} N_{\beta} \sigma_{\alpha\beta}^3 \varepsilon_{\alpha\beta} \int g_{\alpha\beta}^{fs}(\tilde{r}) \tilde{u}^{disp}(\tilde{r}) \tilde{r}^2 d\tilde{r} \quad (\text{C.8})$$

At this point a critical assumption is made, namely that the parameter-weighted average of the two integrals in eq. (C.8) is similar on the left-hand and right-hand side of the equation. This assumption is supported by [183] and leads to the simple relation:

$$\hat{m}^2 \hat{\sigma}^3 \hat{\varepsilon} \approx \sum_{\alpha}^{\text{seg.typ.}} \sum_{\beta}^{\text{seg.typ.}} N_{\alpha} N_{\beta} \sigma_{\alpha\beta}^3 \varepsilon_{\alpha\beta} \quad (\text{C.9})$$

To correct for the assumptions which resulted in eq. (C.9) a second correction parameter ϕ_{ε} is introduced:

$$\hat{m}^2 (\hat{\sigma} \phi_{\sigma})^3 (\hat{\varepsilon} \phi_{\varepsilon}) = \sum_{\alpha}^{\text{seg.typ.}} \sum_{\beta}^{\text{seg.typ.}} N_{\alpha} N_{\beta} \sigma_{\alpha\beta}^3 \varepsilon_{\alpha\beta} \quad (\text{C.10})$$

Equations (C.1),(C.3) and (C.10) can be written in $\{\hat{m}, \hat{\sigma}, \text{ and } \hat{\varepsilon}\}$:

$$\hat{\sigma} = \frac{\left(\frac{\sum_{\alpha}^{\text{seg.typ.}} N_{\alpha} m_{\alpha} \sigma_{\alpha\alpha}^3}{\sum_{\alpha}^{\text{seg.typ.}} N_{\alpha} m_{\alpha}} \right)^{(1/3)}}{\phi_{\sigma}} \quad (\text{C.11})$$

C. Expressing the PC-SAFT model in terms of force field pure parameters

and

$$\hat{\varepsilon} = \frac{\left(\sum_{\alpha}^{\text{seg,typ.}} \sum_{\beta}^{\text{seg,typ.}} N_{\alpha} N_{\beta} \sigma_{\alpha\beta}^3 \varepsilon_{\alpha\beta} \right)}{\hat{m}^2 (\hat{\sigma}_{\sigma})^3 \phi_{\varepsilon}}, \quad (\text{C.12})$$

where \hat{m} and $\hat{\sigma}$ can be inserted from eq. (C.1) and (C.11), respectively. With these expressions the PC-SAFT eos is expressed explicitly in terms of the force field parameters.

D. Influence of the energy bin width on results at low temperatures

D.1. Background to the energy bin width

In this Appendix the influence of the energy bin width ΔE as an input parameter for simulations is analyzed. The bin width ΔE is used for collecting histograms of $H(N,E)$ by counting conditions found between E_j and $E_j + \Delta E$, where the energy is discretized with a spacing of $E_{j+1} = E_j + \Delta E$. Parameter ΔE is one of many parameters used in MC simulations, such as a maximum displacement, number of configurational bias steps, number of equilibration steps and so forth. These parameters should either determine statistical uncertainties (CBMC steps, maximum displacement) or eliminate systematic dependencies on starting conditions (number of equilibration steps). These simulation parameters should, if they are well-chosen, not significantly alter the results when varied in a certain range around the selected setting. For the energy bin width, I had tested variations in ΔE values. I defined ΔE to roughly ensure 200 “observed energy bins” with $E_0 < E_{\min} < E_1$ and $E_{199} < E_{\max} < E_{200}$, with E_{\min} and E_{\max} as the minimal and maximal energy value observed during a simulation.

After completing the study presented in this thesis, however, I realized that for polar species with intramolecular degrees of freedom there is a higher than previously expected dependence of simulation results (from post-processing histogram reweighting) on the choice of parameter ΔE . This appendix illustrates the dependence of calculated phase equilibrium results on varying values of ΔE . At the time of writing this thesis the underlying reason for the observed sensitivity towards ΔE is not fully understood and the analysis of the issue in collaboration with my supervisor J. Gross is ongoing. We currently consider histogram-free methods, like MBAR [106] or a bin-free histogram approach, as proposed by Berg and Harris [184]. Once conclusive insights into a proper choice for ΔE and a method to ensure robust results are obtained, we might summarize and publish these findings in an erratum.

In the following I will show that parameter ΔE affects calculated phase equilibria of low temperatures to an extent that is outside the statistical uncertainty of the simulation method. Fortunately the sensitivity is essentially limited to vapor pressure (and thus

D. Influence of the energy bin width on results at low temperatures

coexisting vapor density) data at low temperatures.

Figures D.1 and D.2 show the relative deviations for the vapor pressure of ethylacetate compared to quasi-experimental data for different values of the energy width input parameter. The different colors denote the different input values of the energy width. Below a reduced temperature of about $T_R < 0.7$ the curves start to deviate.

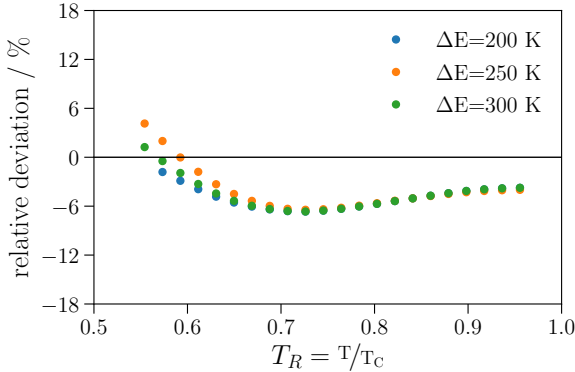


Figure D.1.: Relative deviations of calculated vapor pressure from quasi-experimental data [137] for ethylacetate in the reduced temperature range from $T_R = 0.56 - 0.96$. Blue, orange and green spheres show the relative deviations of simulation results using an energy bin width of $\Delta E = 200$ K, 250 K, and 300 K, respectively

In figure D.1 it seems that this input value has only a small impact on the final result. Upon further increasing the bin width, however, a pronounced dependence of calculated vapor pressure at low temp. is observed, as shown in fig. D.2. However, a look at figure D.2 changes the picture. Increasing the energy width more to $\Delta E = 325$ K and $\Delta E = 350$ K significantly changes the calculated vapor pressure for low temperatures.

D.2. Bin sizes used in this thesis

Table D.1 gives the values of all energy bin widths used in this work, according to the respective substance.

D. Influence of the energy bin width on results at low temperatures

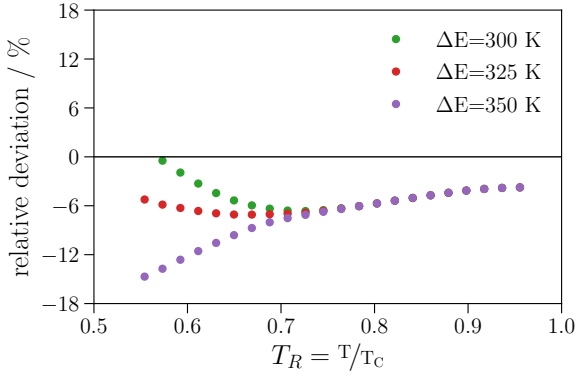


Figure D.2.: Relative deviations of calculated vapor pressure from quasi-experimental data [137] for ethylacetate in the reduced temperature range from $T_R = 0.56 - 0.96$. Green, red and purple spheres show the relative deviations of simulation results using an energy bin width of $\Delta E = 300$ K, 325 K, and 300 K, respectively

Table D.1.: Tabulated energy bin widths used for molecular simulations in this work for each substance.

substance	bin width /K	substance	bin width /K
propane	100	1-undecene	200
butane	100	1-dodecene	200
pentane	100	1-tridecene	200
hexane	100	4-methyl-1-pentene	100
heptane	100	5-methyl-1-hexene	150
octane	100	diethyl-ether	250
nonane	100	ethyl-propyl-ether	250
decane	100	di-n-propyl-ether	250
undecane	200	di-n-butyl-ether	250
dodecane	200	di-n-pentyl-ether	250
hexadecane	200	butanone	250
eicosane	200	2-pentanone	250
tetracosane	200	2-hexanone	250
2-methylpentane	150	2-heptanone	250
3-methylpentane	150	2-octanone	250
2-methylhexane	150	ethyl acetate	250
2-methylheptane	150	propyl acetate	250
1-pentene	200	butyl acetate	250
1-hexene	200	pentyl acetate	250
1-octene	200	ethyl formate	250
1-nonene	200	propyl formate	250
1-decene	200	butyl formate	250

E. Supporting information for chapter 3

This supporting information gives detailed simulation results in tabulated and graphical form.

E.1. Background to reported results

E.1.1. Errors

All given uncertainties were calculated using the 95% confidence interval. The numbers in brackets are the uncertainties of the calculated sample mean from n independent samples $\text{err} = t^*s/\sqrt{n}$, where s is the corrected sample standard deviation

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2},$$

and t^* is the critical value from the student's t-distribution corresponding to the degrees of freedom $\nu = n - 1$. The value of t^* is given in the tables below. If there is no number in brackets, the results are from a single simulation.

E.1.2. Enthalpies of vaporization

The Clausius-Clapeyron equation can readily be evaluated from histogram reweighting techniques. The gradient dp/dT was determined by calculating the gradient $d(\ln p)/d(1/T)$, which is linear, and using

$$\frac{d(\ln p)}{d(1/T)} = -\frac{T^2}{p} \cdot \frac{dp}{dT}. \quad (\text{E.1})$$

The so-obtained results of the enthalpy of vaporization $\Delta^{lv}h$ are less sensitive to finite size simulation boxes as compared to values from $\langle u \rangle^v - \langle u \rangle^l + p^{\text{sat}} (\langle v \rangle^v - \langle v \rangle^l)$.

The detailed description, starting from the Clausius Clapeyron equation, is as follows:

$$\frac{dp}{dT} = \frac{\Delta h_{LV}}{\Delta v_{LV} \cdot T}. \quad (\text{E.2})$$

Applying the chain rule to equation (E.1) and replacing the gradients through central differences

$$\frac{dp}{dT} = -\frac{d(\ln p)}{d(1/T)} \cdot \frac{p}{T^2} \approx \frac{\Delta(\ln p)}{\Delta\left(\frac{1}{T}\right)} \cdot \frac{p}{T^2} \quad (\text{E.3})$$

and inserting it to the Clausius Clapeyron equation

$$\Delta h_{LV} \approx \frac{\Delta(\ln p)}{\Delta\left(\frac{1}{T}\right)} \cdot \frac{p}{T} \cdot \Delta v_{LV} \quad (\text{E.4})$$

With simulation results in form of the equation finally writes as

Entry no.	Temperature	Pressure	Liquid density	Vapor density
$i - 1$	T_{i-1}	p_{i-1}	$\rho_{L,i-1}$	$\rho_{V,i-1}$
i	T_i	p_i	$\rho_{L,i}$	$\rho_{V,i}$
$i + 1$	T_{i+1}	p_{i+1}	$\rho_{L,i+1}$	$\rho_{V,i+1}$

$$\Delta h_{LV} \approx \frac{\ln p_{i+1} - \ln p_{i-1}}{\frac{1}{T_{i+1}} - \frac{1}{T_{i-1}}} \cdot \frac{p_i}{T_i} \cdot \left(\frac{1}{\rho_{V,i}} - \frac{1}{\rho_{L,i}} \right). \quad (\text{E.5})$$

This last equation was used to determine the enthalpies of vaporization in this thesis.

E.1.3. Critical points

Critical values were extrapolated from at least three simulations at different volumes using finite-size scaling techniques[13, 185]. Data for comparison are mostly taken from NIST[186] or DIPPR[137]. ^a We typically consider $V = 22000 \text{ \AA}$, $V = 40000 \text{ \AA}$ and $V = 60000 \text{ \AA}$.

E.2. Results for cyclopentane and cyclohexane

Table E.2.: Critical properties of cyclic alkanes

substance	simulation results			experimental data[186]		
	T_C /K	p_C /bar	ρ_C /kg/m ³	T_C /K	p_C /bar	ρ_C /kg/m ³
cyclopentane	518.5	49.0	273	511.6(3)	45.1(4)	270(3)
cyclohexane	557.5	43.5	274	554(1)	40.7(5)	274(6)

^aData reported by DIPPR/NIST sometimes represent averages of several reported values.

Table E.1.: Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization for cyclic alkanes

cyclopentane				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
300	742.1(6)	1.314(7)	0.459(2)	404.3(3)
340	701.0(7)	4.56(2)	1.752(9)	376.9(5)
380	657(1)	12.02(5)	4.89(2)	344.3(9)
420	606.1(4)	26.8(2)	11.06(4)	303.8(5)
460	543(3)	55(1)	21.6(1)	248(4)
cyclohexane				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
320	746(4)	1.032(5)	0.321(1)	377(2)
360	708(2)	3.53(2)	1.206(3)	353(1)
400	667.3(6)	9.28(2)	3.364(6)	325.7(4)
440	621.8(6)	20.6(1)	7.64(2)	292.5(2)
480	568.0(6)	41.6(7)	15.06(8)	250(2)
500	535(6)	59(2)	20.3(2)	221(5)

E.3. Results for ketones

Table E.3.: Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization for ketones

acetone				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
285	780	0.414	0.167	558
325	734	1.95	0.877	520
365	689	6.28	3.03	479
405	638	16.07	7.98	428
445	577	36.1	17.4	362
485	492	81.1	33.5	263
495	461	103	38.9	226
butanone			4 independent runs	$t^* = 3.182$
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
300	765(3)	0.438(3)	0.1501(9)	476(2)
340	724(4)	1.96(1)	0.743(5)	443(1)
380	681(1)	6.13(2)	2.49(1)	407.0(8)
420	632(2)	15.40(6)	6.44(2)	364(1)
460	575(2)	34.0(1)	13.91(2)	310(1)
500	497(1)	73.3(5)	26.54(6)	231.1(4)
2-pentanone			5 independent runs	$t^* = 2.776$
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
320	755(3)	0.408(1)	0.1239(3)	439(2)
360	716.5(5)	1.783(4)	0.595(1)	407.7(2)
400	676.8(5)	5.472(9)	1.970(3)	375.5(3)
440	630.9(6)	13.54(1)	5.043(5)	337.8(4)
480	579(1)	29.61(7)	10.86(1)	291.8(6)
520	511.1(6)	61.3(2)	20.66(4)	229.3(5)
535	476.8(9)	82.1(4)	25.67(5)	195.7(6)

E. Supporting information for chapter 3

3-pentanone 1 run				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
320	749	0.461	0.140	430
360	712	1.95	0.651	403
400	671	5.93	2.11	371
440	625	14.45	5.33	332
480	572	31.3	11.33	286
520	501	65.1	21.4	222
535	464	88.6	26.5	187
2-hexanone 4 independent runs $t^* = 3.182$				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
340	747(2)	0.424(7)	0.119(2)	403(2)
380	709(5)	1.66(2)	0.518(5)	378(3)
420	672(2)	4.96(3)	1.63(1)	351.8(6)
460	629(2)	12.20(5)	4.12(2)	319(1)
500	580(2)	26.5(1)	8.84(2)	278(1)
540	519(1)	54.0(3)	16.83(5)	224.3(8)
560	478(2)	78.1(8)	22.43(8)	187(2)
2-heptanone 4 independent runs $t^* = 3.182$				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
340	755(2)	0.185(1)	0.0455(3)	395(1)
380	721(3)	0.874(5)	0.238(1)	371(2)
420	685(3)	2.90(1)	0.850(3)	346(1)
460	646(1)	7.64(6)	2.34(1)	317(1)
500	602(1)	17.4(1)	5.37(3)	283.2(8)
540	550(2)	36.1(3)	10.75(5)	241(1)
580	480(3)	71(1)	19.4(1)	183(3)
2-octanone 4 independent runs $t^* = 3.182$				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
360	744(5)	0.205(1)	0.0475(2)	376(2)
400	712(4)	0.912(9)	0.232(2)	353(2)
440	677(2)	2.93(2)	0.799(4)	329(1)
480	639(2)	7.55(3)	2.153(5)	301.3(7)
520	596.2(5)	16.94(7)	4.86(1)	269.4(4)
560	546(1)	35.1(1)	9.67(2)	229.3(8)
600	479(3)	69.6(8)	17.42(5)	175(2)

Table E.4.: Critical properties of simulated ketones

substance	simulation results			experimental data 137 , 186		
	T_C /K	p_C /bar	ρ_C /kg/m ³	T_C /K	p_C /bar	ρ_C /kg/m ³
acetone	519.9	55.7	269	508.15	47.6	277.9
butanone	537.9	45.3	266	535(2)	42(2)	252 - 270
2-pentanone	569.6	41.1	260	561.1(2)	37.9(20)	286(8)
3-pentanone	562.4	38.4	270	561.5(2)	37.3(1)	256(30)
2-hexanone	596.4	36.5	257	586.8(6)	33.2(8)	267(3)
2-heptanone	620.7	32.7	255	611.5(4)	34.40(4)	261(3)
2-octanone	644.5	31.3	251	632.7(2)	26.4	258(3)

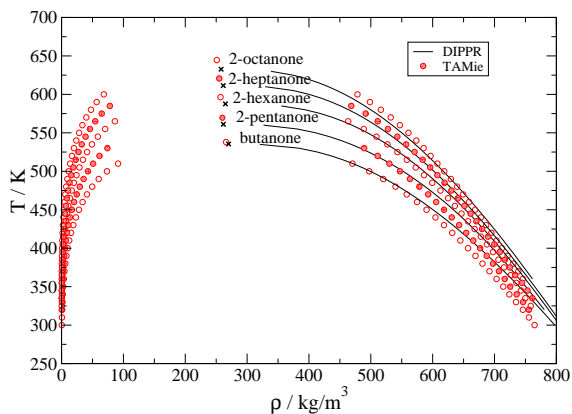


Figure E.1.: Phase diagram of simulated ketones. Lines are quasi-experimental data[137]. Symbols represent simulation results using the TAMie force field.

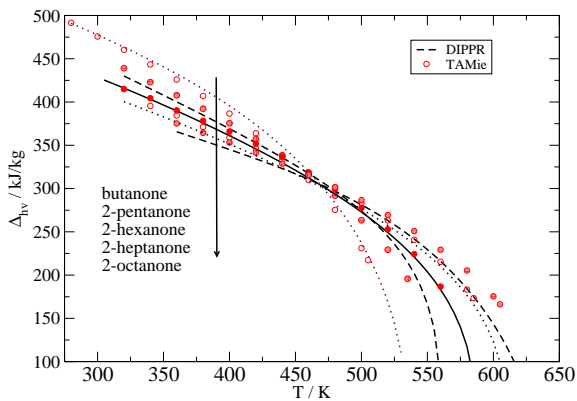


Figure E.2.: Ketones: Enthalpies of vaporization. Red symbols result from simulation with the TAMie force field. Dashed, dotted and solid lines correspond to DIPPR correlations[137].

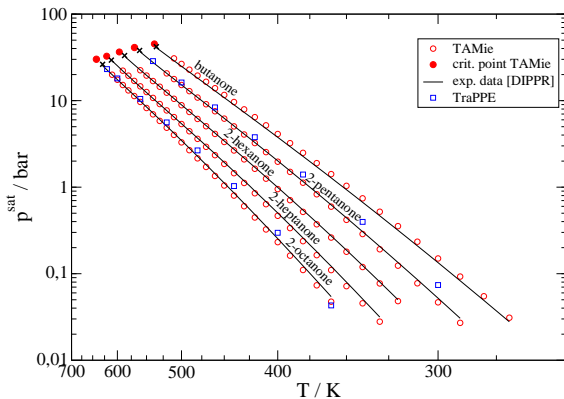


Figure E.3.: Vapor pressures of ketones compared to results of the TraPPE[53] united atom force field and quasi-experimental data from DIPPR[137].

E.4. Results for aldehydes

Table E.5.: Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization for aldehydes

propanal		1 run			
T / K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / bar	$\Delta^{\text{lv}}h$ / kJ/kg	
280	802	0.52	0.205	533	
320	756	2.39	1.05	597	
360	708	7.56	3.59	456	
400	655	19.0	9.29	405	
440	590	42.4	20.0	339	
480	497	94.7	38.0	239	
butanal		1 run			
T / K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / bar	$\Delta^{\text{lv}}h$ / kJ/kg	
300	786	0.43	0.147	482	
340	746	1.99	0.748	450	
380	702	6.22	2.54	413	
420	652	15.4	6.53	370	
460	595	33.6	14.0	317	
500	520	70.4	26.5	246	
520	464	108	35.3	189	
pentanal		4 independent runs		$t^* = 3.182$	
T / K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / bar	$\Delta^{\text{lv}}h$ / kJ/kg	
300	792	0.16	0.047	463	
340	755(2)	0.868(2)	0.281(3)	433(3)	
380	715(1)	3.06(1)	1.081(4)	400(4)	
420	673(1)	8.26(4)	3.082(4)	366(4)	

E. Supporting information for chapter 3

460	625(1)	18.8(1)	7.15(5)	326(3)
500	568(1)	39.1(3)	14.32(5)	277(3)
530	513(3)	67(2)	22.5(1)	227(1)

Table E.6.: Critical properties of simulated aldehydes

substance	simulation results			experimental data 137 , 186		
	T_C /K	p_C /bar	ρ_C /kg/m ³	T_C /K	p_C /bar	ρ_C /kg/m ³
propanal	508.8	56.8	284	504.4(47)	52.6(6)	285(5)
butanal	542.4	47.4	279	537.2(8)	43.2(10)	280(5)
pentanal	576.0	41.3	273	567(2)	39.7(10)	275(5)
hexanal	604.4	36.8	268	592(3)	34.6(5)	265(5)
octanal	653.0	31.6	256	639(1)	29.6(5)	263(4)

E. Supporting information for chapter 3

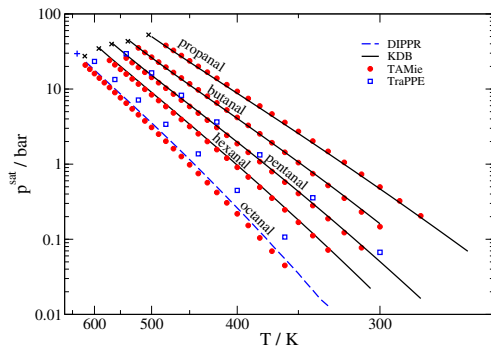


Figure E.4.: Vapor pressures of aldehydes compared to results of the TraPPE[53] united atom force field and quasi-experimental data from the KDB[138] and DIPPR[137].

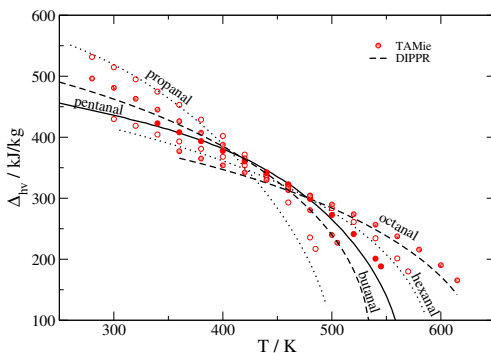


Figure E.5.: Aldehydes: Enthalpies of vaporization. Red symbols result from simulation with the TAMie force field. Dashed, dotted and solid lines correspond to correlations of the DIPPR[137] data base.

Table E.7.: Consistency check using GEMC with CASSANDRA for 2-pentanone. 10 Mio equilibration steps + 10 Mio production steps in Gibbs Ensemble. Errors are calculated using the reblocking method of Flyvbjerg and Petersen[2].

T /K	CASSANDRA			Our results		
	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar
400	675(2)	5.4(3)	1.97(9)	676.8(5)	5.472(9)	1.970(3)
480	578(1)	28.8(8)	10.7(1)	579(1)	29.61(7)	10.86(1)

E.5. Confirming simulation results with other simulation software

The simulation results reported in this work was obtained with a molecular simulation code developed in our group. In order to demonstrate that the results are meaningful, we perform calculations with CASSANDRA[87]. We choose 2-pentanone as an example, because this substance possesses a “branching site” as well as partial charges.

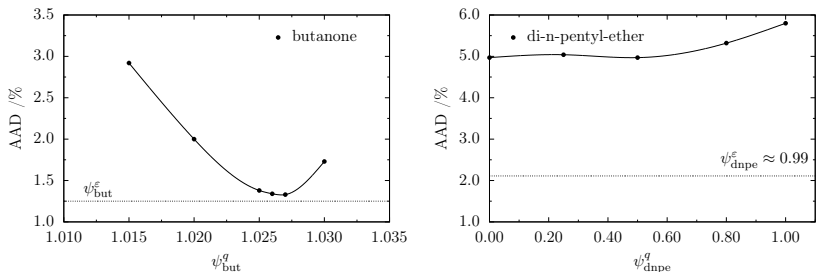
F. Supporting information for chapter 4

This supporting information gives simulation details and results in tabulated and graphical form.

F.1. Alternative scaling schemes

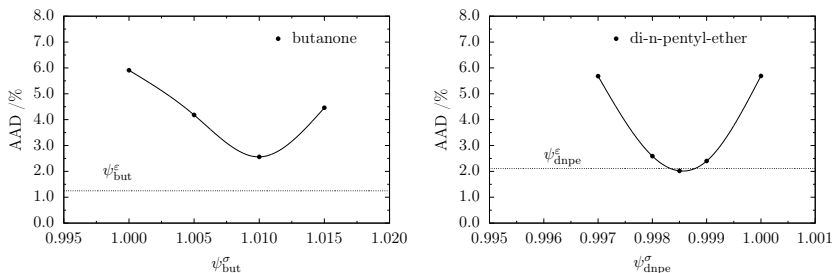
We compare the approach of scaling all ε_i parameters with two alternative scaling schemes, namely scaling the Mie size parameter according to $\sigma_{ii}^{\text{ind}} = \psi_A^\sigma \cdot \sigma_{ii}$ and scaling the electrostatic potential as $q_i^{\text{ind}} = \psi_A^q \cdot q_i$. The comparison is made for 2 representative substances (butanone and di-n-pentyl-ether). Figures F.1a and F.1b confirm that scaling the Mie energy parameter according to $\varepsilon_{ii}^{\text{ind}} = \psi_A \cdot \varepsilon_{ii}$ (as finally chosen) is indeed a promising approach, because it is easy to apply and works well for all considered substances. Figure F.1a shows that scaling the electrostatic potential with $\psi_{\text{but}}^q \approx 1.26$ leads to a similar improvement for butanone as scaling the Mie energy parameter with $\psi_{\text{but}} \approx 1.12$. However, the picture is entirely different for di-n-pentyl-ether. Figure F.1b shows that even turning of the electrostatic interactions completely – corresponding to $\psi_{\text{dnpe}}^q = 0.0$ does not lead to a significant improvement of the absolute average deviation. The error can only be reduced from about 5.8 % to 5.0 %, while when applying $\varepsilon_{ii}^{\text{ind}} = \psi_A \cdot \varepsilon_{ii}$ the error is about 2.1 % for $\psi_{\text{dnpe}} \approx 0.99$. Scaling the Mie size parameters worked well for di-n-pentyl-ether with AADs similar or even slightly better than when scaling the Mie energy parameters. But, this behavior did not hold for butanone, where improvements could be achieved but stood behind the improvements as when scaling σ . We therefore conclude that scaling the Mie energy parameter seems to be the easiest and most promising concept if a transferable force field should be individualized.

F. Supporting information for chapter 4



(a) Scaling behavior of butanone when scaling the Mie energy parameters. (b) Scaling behavior of di-n-pentyl-ether when scaling the Mie energy parameters.

Figure F.1.: Scaling the Mie energy parameters or the electrostatic potential lead to similar results for butanone with AADs about 1.25%. However, for di-n-pentyl-ether scaling the Mie energy parameters does not lead to any significant improvement, even when turning off the electrostatic interactions completely. Filled spheres denote results from simulation with different point charges and the dashed line denotes the result from iTAMie. The solid line is a guide for the eye.



(a) Scaling behavior of butanone when scaling the Mie size parameters. (b) Scaling behavior of di-n-pentyl-ether when scaling the Mie size parameters.

Figure F.2.: Scaling the Mie size parameters σ_i or the Mie energy parameters ϵ_i lead to similar results for di-n-pentyl-ether with AADs about 2.1% with a slightly less AAD when scaling σ_i . For butanone an improvement is visible when scaling σ_i . But, the minimum has an AAD about 2.5% while the errors could be reduced to 1.25% when scaling the Mie energy parameters ϵ_i .

F.2. Simulation Details

Table G.1 gives simulation details for every substance we have individualized. The table gives information about the number of equilibration moves, production moves, the volume of the simulation box the width of each window and the maximum number of molecules. The maximum number of particles N^{\max} divided by the width of a window ΔN gives the number of used windows. Two windows have been simulated on one CPU core. For propane for example, the used number of CPU cores was $(432/12)/2 = 18$.

Configurational Bias Monte Carlo

As mentioned in the main article, the sampling of insertion, deletion, or regrowth steps was enhanced using the configurational biasing (CB) method. Since for the lowest densities the acceptance probability is reasonably high, we perform only one CB step for densities up to a dimensionless density of $\rho^* \leq 1/2.7$, with $\rho^* = \rho/\rho_{\max}$, where ρ_{\max} again is the highest density for which we perform simulations. Table G.2 shows the dimensionless densities and the according CB steps.

F.3. Correlation function parameters for the newly fit vapor pressure correlation of ethyl-propyl-ether

The vapor pressure correlation function of ethyl-propyl-ether was fit to experimental data of the DIPPR data base. The equation is as follows

$$\ln(p^{\text{sat}})/\text{Pa} = A + \frac{B}{T/\text{K}} + C \cdot \ln(T/\text{K}) + D \cdot (T/\text{K})^E \quad (\text{F.1})$$

with

F. Supporting information for chapter 4

Table F.1.: simulation details

substance	N^{equil} /mio.	N^{prod} /mio.	Volume / \AA^3	ΔN	N^{max}
propane	5	20	50000	12	432
n-butane	5	20	60000	12	408
n-pentane	20	25	60000	6	336
n-hexane	5	20	80000	12	384
n-heptane	5	20	80000	12	336
n-octane	5	20	80000	10	300
n-nonane	5	20	80000	9	270
n-decane	7	30	80000	5	240
n-undecane	8	25	100000	5	280
n-dodecane	7	30	100000	8	256
n-tridecane	8	25	120000	5	289
n-hexadecane	7	25	100000	6	192
n-icosane	7	25	130000	4	200
n-tetracosane	7	25	160000	4	208
2-methylpentane	7	25	65000	8	320
3-methylpentane	7	25	65000	8	320
2-methylhexane	7	25	80000	6	336
2-methylheptane	7	25	75000	5	280
1-pentene	8	30	60000	9	360
1-hexene	5	25	60000	8	304
1-octene	5	25	80000	5	310
1-nonene	8	25	81000	5	280
1-decene	8	25	90000	5	280
1-undecene	8	25	105000	5	300
1-dodecene	8	25	115000	5	300
1-tridecene	8	25	125000	5	300
4-methyl-1-pentene	7	25	60000	8	304
5-methyl-1-hexene	7	25	70000	6	312
diethyl-ether	5	20	49000	5	310
ethyl-propyl-ether	9	21	60000	8	320
dipropyl-ether	5	20	60000	5	270
dibutyl-ether	8	25	80000	9	288
dipentyl-ether	8	30	103000	5	300
butanone	5	25	47000	10	320
2-pentanone	5	25	56000	10	320
2-hexanone	6	25	65000	10	320
2-heptanone	7	25	71000	10	300
2-octanone	7	25	80000	10	300

Table F.2.: Number of CB steps for different dimensionless densities

ρ^*	CB steps per bead
$< 1/2.7$	1
$\geq 1/2.7 \vee < 1/1.9$	2
$\geq 1/1.9 \vee < 1/1.6$	3
$\geq 1/1.6 \vee < 1/1.3$	4
$\geq 1/1.3 \vee < 1/1.16$	6
$\geq 1/1.16$	8

Table F.3.: Parameters for the Clausius-Clapeyron like fit of the vapor pressure for ethyl-propyl-ether

A	63.1869
B	-5655.14
C	-5.99446
D	$2.35 \cdot 10^{-17}$
E	6

F.4. Enthalpies of vaporization

The Clausius-Clapeyron equation can readily be evaluated from histogram reweighting techniques. The gradient dp/dT was determined by calculating the gradient $d(\ln p)/d(1/T)$ and using

$$\frac{d(\ln p)}{d(1/T)} = -\frac{T^2}{p} \cdot \frac{dp}{dT}.$$

The so-obtained results of the enthalpy of vaporization $\Delta^{lv}h$ are less sensitive to finite size simulation boxes as compared to values from $\langle u \rangle^v - \langle u \rangle^l + p^{\text{sat}} (\langle v \rangle^v - \langle v \rangle^l)$.

F.5. Graphical results

Here we show the absolute average deviation (AAD) for every substance in three figures.

F. Supporting information for chapter 4

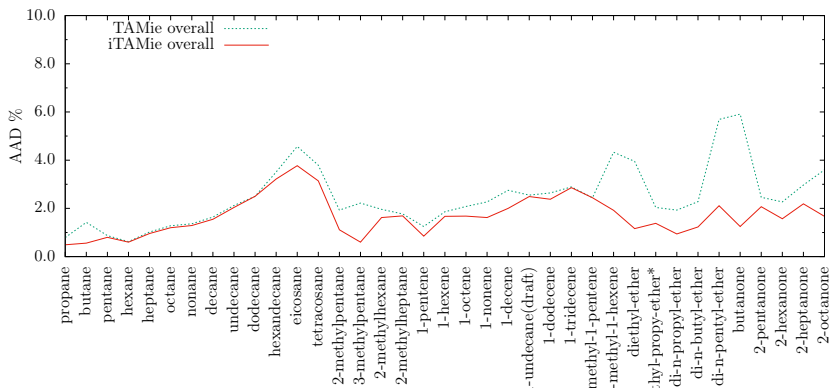


Figure F.3.: AAD combined for all individualized substances

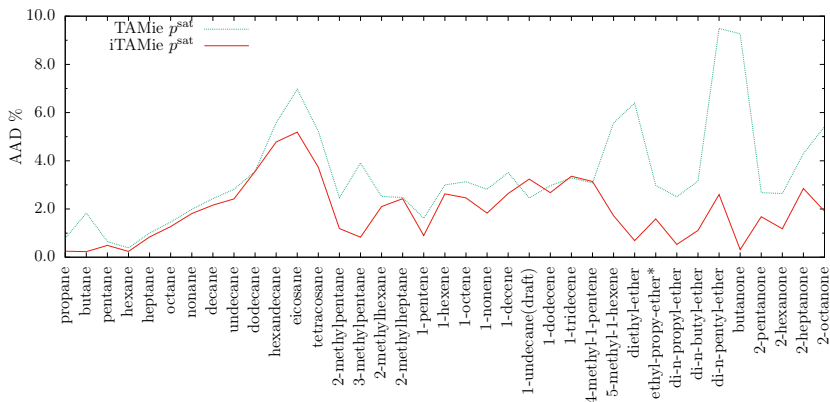


Figure F.4.: AAD of the vapor pressure for all individualized substances

F. Supporting information for chapter 4

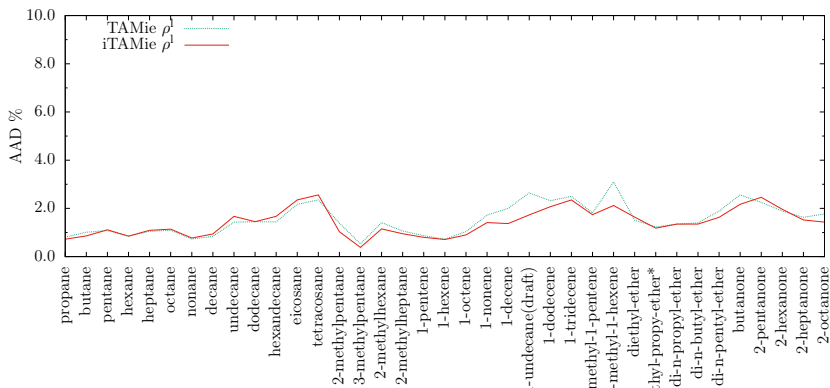


Figure F.5.: AAD of the liquid density for all individualized substances

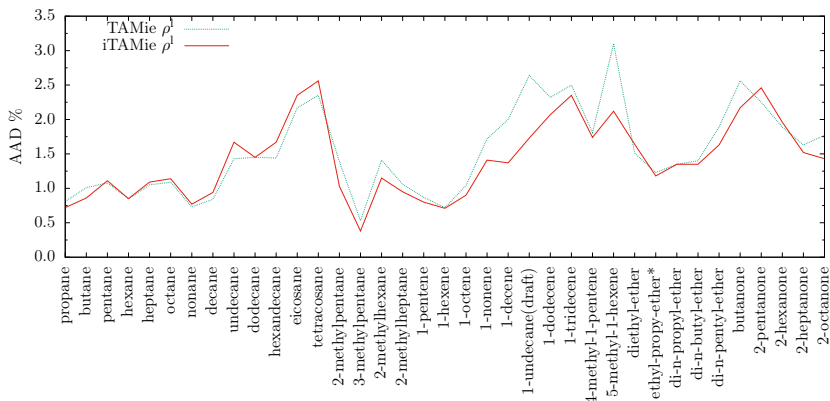


Figure F.6.: AAD of the liquid density for all individualized substances

F. Supporting information for chapter 4

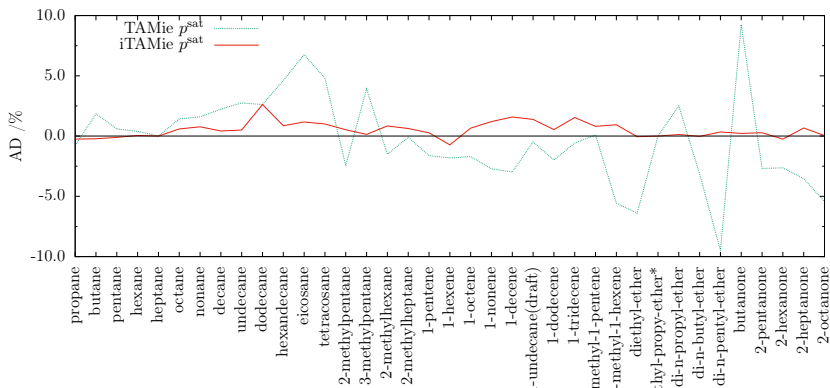


Figure F.7.: AD of the vapor pressure for all individualized substances

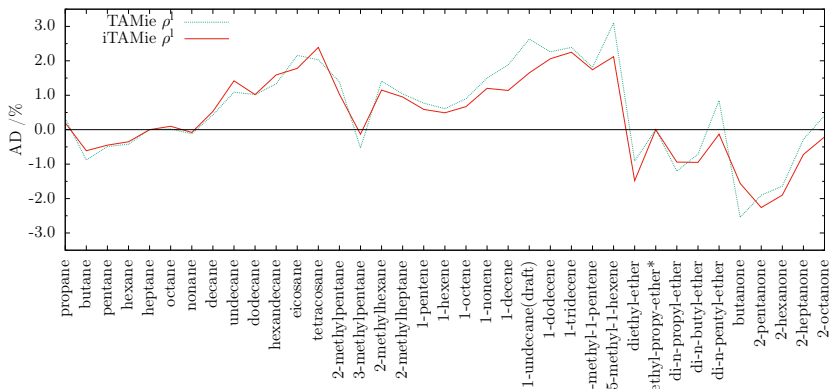


Figure F.8.: AD of the liquid density for all individualized substances

F. Supporting information for chapter 4

Table F.4.: Optimization parameters ψ_A and AAD of TAMie and iTAMie compared to DIPPR correlations for different substances. Ethyl-propyl-ether* has been adjusted to quasi-experimental data of a newly fit correlation function.

substance	AAD TAMie /%	AAD iTAMie /%	AD p^{sat} TAMie /%	AD p^{sat} iTAMie /%	AD ρ^{l} TAMie /%	AD ρ^{l} iTAMie /%
propane	0.79	0.51	-0.78	0.45	0.33	-0.07
butane	1.42	0.54	1.83	-0.22	-0.88	-0.61
pentane	0.87	0.80	0.60	-0.11	-0.49	-0.45
hexane	0.62	0.55	0.39	0.05	-0.42	-0.35
heptane	1.02	0.96	0	0	0	0
octane	1.28	1.20	1.42	0.59	0.02	0.10
nonane	1.36	1.29	1.60	0.77	-0.12	-0.08
decane	1.64	1.55	2.23	0.42	0.43	0.53
undecane	2.12	2.04	2.75	0.50	1.09	1.42
dodecane	2.5	2.5	2.62	2.62	1.02	1.02
hexadecane	3.51	3.22	4.61	0.86	1.33	1.59
eicosane	4.57	3.77	6.76	1.17	2.16	1.78
tetracosane	3.79	3.14	4.80	1.00	2.03	2.39
2-methylpentane	1.93	1.11	-2.45	0.53	1.39	1.03
3-methylpentane	2.22	0.60	3.91	0.14	-0.53	-0.13
2-methylhexane	1.93	1.62	-1.50	0.84	1.41	1.15
2-methylheptane	1.77	1.69	-0.10	0.63	1.04	0.95
1-pentene	1.25	0.85	-1.62	0.27	0.77	0.59
1-hexene	1.86	1.67	-1.82	-0.73	0.61	0.49
1-octene	2.08	1.68	-1.70	0.66	0.90	0.67
1-nonene	2.27	1.62	-2.72	1.20	1.50	1.20
1-decene	2.75	2.00	-2.98	1.58	1.88	1.14
1-undecene	2.55	2.49	-0.49	1.38	2.63	1.65
1-dodecene	2.64	2.38	-1.99	0.54	2.26	2.06
1-tridecene	2.89	2.86	-0.58	1.54	2.39	2.25
4-methyl-1-pentene	2.44	2.44	0.10	0.81	1.81	1.74
5-methyl-1-hexene	4.33	1.92	-5.57	0.94	3.10	2.12
diethyl-ether	3.95	1.16	-6.39	-0.05	-0.91	-1.48
ethyl-propyl-ether*	2.09	1.29	0	0	0	0
di-n-propyl-ether	1.93	0.94	2.51	0.13	-1.20	-0.94
di-n-butyl-ether	2.28	1.23	-3.15	-0.02	-0.71	-0.95
di-n-pentyl-ether	5.69	2.11	-9.49	0.34	0.85	-0.13
butanone	5.91	1.25	9.27	0.22	-2.54	-1.56
2-pentanone	2.46	2.07	-2.68	0.28	-1.90	-2.26
2-hexanone	2.27	1.57	-2.64	-0.25	-1.64	-1.90
2-heptanone	2.97	2.19	-3.54	0.67	-0.28	-0.72
2-octanone	3.60	1.67	-5.42	0.04	0.41	-0.21

F. Supporting information for chapter 4

Table F.5.: Alkanes: Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization. The last line for each substance gives the estimated critical properties.

propane				
l run				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ /kJ/kg
200	614	0.545	0.204	458
240	567	3.44	1.49	417
280	518	12.5	5.84	370
320	456	34.7	16	301
355	375	83	32.4	201
373	225	225	44.4	0
n-butane				
l run				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ /kJ/kg
240	629	0.71	0.241	417
280	586	3.47	1.33	381
320	541	11.1	4.56	342
360	487	28.4	11.7	289
400	412	68.1	24.7	208
410	385	87.5	29.3	177
429	233	233	39.6	0
n-pentane				
l run				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ /kJ/kg
280	630	1.05	0.333	379
320	591	4.13	1.44	350
360	549	11.8	4.35	315
400	498	28.2	10.4	269
440	430	62.8	21	202
455	391	88.6	26.5	162
475	236	236	35.7	0
n-hexane				
l run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg
280	668	0.32	0.0857	383
320	629	1.6	0.48	353
360	590	5.3	1.73	324
400	547	13.8	4.66	290
440	498	31.4	10.4	247
480	428	67.5	20.1	183
495	389	93.5	25.1	146
514	237	237	32.8	0
n-heptane				
l run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg
280	685	0.1	0.0233	339
320	653	0.632	0.166	350
360	619	2.47	0.712	328
400	582	7.18	2.19	300
440	539	17.3	5.37	267
480	488	37.1	11.2	224
520	415	79.2	20.8	158
548	239	239	30.5	0
n-octane				
l run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg
320	673	0.248	0.0574	350
360	642	1.15	0.296	329
400	607	3.77	1.04	305
440	569	9.81	2.82	277
480	525	22.4	6.36	243

F. Supporting information for chapter 4

520	471	46.8	12.6	198
550	410	82.6	19.7	145
578	239	239	28.4	0

F. Supporting information for chapter 4

n-nonane				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^l v h$ / kJ/kg
280	705	0.00969	0.00176	357
320	689	0.0985	0.0204	348
360	658	0.547	0.126	328
400	626	2.02	0.507	307
440	592	5.74	1.52	284
480	552	13.7	3.71	255
520	505	29.8	7.78	218
560	445	60.1	14.6	169
575	410	79.8	18	141
605	238	238	26.6	0
n-decane				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^l v h$ / kJ/kg
360	674	0.257	0.0538	329
400	642	1.09	0.25	309
440	610	3.41	0.834	287
480	573	8.73	2.2	262
520	531	19.3	4.91	232
560	483	38.3	9.55	195
595	418	71.3	15.7	145
629	236	236	24.1	0
n-undecane				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^l v h$ / kJ/kg
360	684	0.121	0.0231	327
400	657	0.59	0.124	310
440	626	2.04	0.461	290
480	590	5.59	1.32	267
520	553	13.1	3.14	241
560	508	26.6	6.44	209
600	452	52.2	11.7	166
620	409	75.4	15.4	135
651	234	234	22.7	0
n-tridecane				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^l v h$ / kJ/kg
400	675	0.177	0.0318	308
440	647	0.752	0.147	291
480	615	2.39	0.497	271
520	581	6.23	1.34	249
560	544	14.1	3.07	224
600	499	27.9	6.11	194
640	443	54.7	10.9	153
660	398	77.8	14.2	122
690	232	232	20.3	0
n-hexadecane				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^l v h$ / kJ/kg
440	670	0.171	0.0274	291
480	640	0.69	0.12	274
520	610	2.15	0.393	256
560	577	5.54	1.05	235
600	542	12.5	2.39	213
640	499	24.9	4.78	185
680	446	48.3	8.62	149
700	405	67.8	11.2	121
738	228	228	17.5	0
n-eicosane				
1 run				

F. Supporting information for chapter 4

T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg
440	677	0.0243	0.00314	282
480	659	0.134	0.0189	272
520	634	0.538	0.0812	258
560	606	1.69	0.269	242
600	573	4.44	0.728	223
640	539	10.4	1.7	201
680	498	21.3	3.49	176
720	450	38.8	6.36	148
750	395	61.4	9.34	115
789	218	218	14.6	0
n-tetracosane	1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg
480	675	0.0281	0.0033	229
520	649	0.147	0.0187	256
560	619	0.562	0.076	240
600	593	1.72	0.244	226
640	563	4.41	0.648	209
680	528	9.65	1.47	190
720	489	18.8	2.91	168
760	444	36.2	5.22	139
790	387	58.7	7.74	108
827	215	215	12	0

F. Supporting information for chapter 4

2-methylpentane				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat}/Δ bar	$\Delta^{lv}h$ / kJ/kg
280	668	0.464	0.125	361
320	632	2.15	0.645	337
360	593	6.86	2.21	309
400	545	17.5	5.77	273
440	491	39.1	12.5	227
480	407	89.2	23.8	151
501	244	244	32.5	0
3-methylpentane				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat}/Δ bar	$\Delta^{lv}h$ / kJ/kg
280	675	0.411	0.11	366
320	638	1.94	0.583	340
360	598	6.22	2.02	312
400	554	15.7	5.27	279
440	501	35	11.4	235
480	426	77.7	21.9	166
485	412	87.4	23.6	155
508	245	245	32.8	0
2-methylhexane				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat}/Δ bar	$\Delta^{lv}h$ / kJ/kg
320	660	0.857	0.223	341
360	623	3.24	0.924	315
400	583	9.15	2.75	287
440	536	21.7	6.58	251
480	480	46.1	13.4	205
510	418	83.6	21.2	151
535	244	244	30.1	0
2-methylheptane				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat}/Δ bar	$\Delta^{lv}h$ / kJ/kg
320	679	0.337	0.0779	340
360	645	1.5	0.385	318
400	608	4.77	1.31	294
440	568	12.2	3.45	265
480	521	27.3	7.61	229
520	460	55.9	14.7	180
540	416	83	19.6	144
567	243	243	28.2	0

Table F.6.: Olefins: Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization. The last line for each substance gives the estimated critical properties.

1-pentene				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat}/Δ bar	$\Delta^{lv}h$ / kJ/kg
280	653	1.3	0.418	380
320	613	4.94	1.77	349
360	567	13.8	5.19	313
400	512	32.7	12.1	265
440	437	73.5	24.3	192
450	409	92.7	28.4	164
470	246	246	38.3	0
1-hexene				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat}/Δ bar	$\Delta^{lv}h$ / kJ/kg
280	682	0.395	0.109	374

F. Supporting information for chapter 4

320	648	1.85	0.573	352
360	608	6.01	2	324
400	563	15.5	5.29	289
440	511	34.5	11.6	245
480	438	74.8	22.3	178
485	425	83.4	24	167
512	248	248	35	0
1-octene	1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^v h$ / kJ/kg
360	658	1.3	0.339	329
400	621	4.18	1.17	305
440	582	10.8	3.12	277
480	537	24	6.97	242
520	481	48.5	13.5	198
550	418	84.9	20.8	146
577	245	245	29.4	0
1-nonene	1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^v h$ / kJ/kg
360	675	0.624	0.146	328
400	641	2.26	0.576	307
440	605	6.32	1.7	283
480	564	15.0	4.08	254
520	517	31.8	8.48	217
560	453	62.6	15.6	168
575	418	83.4	19.1	141
602	244	244	27	0
1-decene	1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^v h$ / kJ/kg
360	686	0.297	0.0629	327
400	655	1.23	0.285	308
440	622	3.77	0.933	286
480	583	9.50	2.42	260
520	541	21.1	5.35	229
560	488	42.2	10.4	190
595	423	77.0	17.0	142
626	242	242	25.1	0
1-undecene	1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^v h$ / kJ/kg
360	699	0.14	0.027	327
400	668	0.665	0.141	308
440	635	2.25	0.513	288
480	602	6.07	1.45	266
520	562	14.2	3.41	239
560	517	29.5	7.01	206
600	457	56.2	12.8	164
649	239	239	23.7	0
1-dodecene	1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^v h$ / kJ/kg
400	680	0.358	0.0701	309
440	647	1.34	0.285	290
480	614	3.89	0.872	269
520	578	9.51	2.18	245
560	536	20.2	4.69	217
600	486	39.7	8.87	181
640	410	79	15.4	127
670	237	237	22.2	0
1-tridecene	1 run			

F. Supporting information for chapter 4

T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^v h$ / kJ/kg
400	685	0.195	0.0353	307
440	657	0.811	0.16	291
480	626	2.55	0.534	272
520	590	6.57	1.43	249
560	554	14.7	3.24	225
600	509	28.8	6.4	195
640	451	56.3	11.4	153
690	234	234	21.1	0
4-methyl-1-pentene	1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^v h$ / kJ/kg
280	685	0.574	0.157	361
320	647	2.58	0.789	336
360	605	8.06	2.64	306
400	556	20.2	6.75	269
440	496	44.9	14.4	220
475	419	93.7	25.2	150
496	252	252	34.3	0
5-methyl-1-hexene	1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^v h$ / kJ/kg
280	711	0.171	0.0404	330
320	677	0.991	0.263	339
360	639	3.64	1.06	314
400	598	10.1	3.09	286
440	549	23.7	7.26	249
480	490	50.4	14.7	202
510	424	91.4	23.1	146
533	252	252	31.7	0

F. Supporting information for chapter 4

Table F.7.: Ether and Ketones: Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization. The last line for each substance gives the critical properties.

diethyl-ether		1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg	
240	756	0.151	0.0402	369	
280	712	1.12	0.34	378	
320	668	4.59	1.55	346	
360	620	13.3	4.76	309	
400	559	32.4	11.5	262	
440	479	74.5	23.4	192	
472	266	266	38.6	0	
ethyl-propyl-ether		1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg	
280	731	0.396	0.105	366	
320	692	1.93	0.572	343	
360	647	6.46	2.04	314	
400	600	16.9	5.51	280	
440	540	38.5	12.2	234	
480	455	86.4	23.7	163	
507	263	263	35.2	0	
di-n-propyl-ether		1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg	
280	743	0.137	0.031	362	
320	706	0.843	0.216	339	
360	666	3.25	0.912	315	
400	623	9.34	2.76	286	
440	575	22.6	6.7	251	
480	514	48.3	13.8	205	
510	449	87.4	22	152	
538	260	260	32.3	0	
di-n-butyl-ether		1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg	
320	733	0.142	0.0288	335	
360	695	0.756	0.171	313	
400	661	2.72	0.667	293	
440	621	7.57	1.95	268	
480	577	18	4.65	238	
520	524	38.8	9.65	200	
560	449	79.1	17.8	145	
594	254	254	27.8	0	
di-n-pentyl-ether		1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg	
320	746	0.0236	0.00396	332	
360	718	0.181	0.034	315	
400	683	0.838	0.173	295	
440	649	2.79	0.618	275	
480	611	7.46	1.72	251	
520	569	17.3	3.98	223	
560	518	36	8.09	189	
600	451	72.2	14.7	141	
638	248	248	23.9	0	
butanone		1 run			
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg	
280	786	0.15	0.0481	497	
320	747	0.875	0.317	465	

F. Supporting information for chapter 4

360	707	3.25	1.29	433
400	662	9.08	3.79	394
440	611	21.3	8.97	348
480	548	45.5	18.2	286
515	471	90.9	31	207
542	276	276	45	0
2-pentanone	1 run			
T /K	$\rho^l / \text{kg/m}^3$	$\rho^v / \text{kg/m}^3$	$p^{\text{sat}} / \Delta \text{ bar}$	$\Delta^{lv}h / \text{kJ/kg}$
240	807	0.00372	0.000862	469
280	781	0.057	0.0153	462
320	759	0.424	0.129	441
360	717	1.84	0.615	407
400	674	5.63	2.02	373
440	630	13.9	5.16	336
480	577	30.4	11.1	289
520	509	62.9	21.1	226
565	272	272	39.2	0

F. Supporting information for chapter 4

2-hexanone				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg
320	763	0.189	0.0501	411
360	725	0.903	0.269	389
400	692	3.04	0.974	364
440	651	8.16	2.72	334
480	603	18.7	6.29	297
520	549	39.1	12.6	250
560	472	80.5	22.9	182
593	268	268	35.2	0
2-heptanone				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg
360	740	0.451	0.117	383
400	701	1.74	0.493	356
440	664	5.05	1.52	329
480	623	12.2	3.76	298
520	575	26.3	7.97	260
560	515	53.1	15.1	211
585	463	83.2	21.5	168
615	266	266	31.6	0
2-octanone				
1 run				
T /K	ρ^l / kg/m ³	ρ^v / kg/m ³	p^{sat} / Δ bar	$\Delta^{lv}h$ / kJ/kg
360	744	0.222	0.0513	374
400	711	0.982	0.249	351
440	675	3.13	0.851	325
480	636	8	2.27	298
520	593	17.9	5.1	265
560	541	37.2	10.1	224
600	470	74.3	18.2	168
610	444	90.2	20.8	146
639	261	261	29.9	0

G. Supporting information for chapter 5

This supporting information gives detailed simulation results in tabulated and graphical form.

G.1. Background to reported results

G.1.1. Enthalpies of vaporization

The Clausius-Clapeyron equation can readily be evaluated from histogram reweighting techniques. The gradient dp/dT was determined by calculating the gradient $d(\ln p)/d(1/T)$ and using

$$\frac{d(\ln p)}{d(1/T)} = -\frac{T^2}{p} \cdot \frac{dp}{dT}.$$

The so-obtained results of the enthalpy of vaporization $\Delta^{lv}h$ are less sensitive to finite size simulation boxes as compared to values from $\langle u \rangle^v - \langle u \rangle^l + p^{\text{sat}} (\langle v \rangle^v - \langle v \rangle^l)$.

G.1.2. Critical points

The critical values given here are size-dependent. No extrapolations with finite-size scaling techniques were examined, since critical data was not in the focus of our force field optimization. However, the estimated critical points are in good agreement to experimental data even if they might get slightly worse when extrapolating to infinite volume.

G.2. Simulation Details

Table G.1 gives simulation details for every substance we have individualized. The table gives information about the number of equilibration moves, production moves, the volume of the simulation box the width of each window and the maximum number of molecules. The maximum number of particles N^{max} divided by the width of a window ΔN gives the number of used windows. Two windows have been simulated on one CPU core. For ethyl acetate for example, the used number of CPU cores was $(320/8)/2 = 20$.

Table G.1.: simulation details

substance	N^{equil} /mio.	N^{prod} /mio.	Volume / \AA^3	ΔN	N^{max}
ethyl formate	7	25	50000	6	384
propyl formate	6	25	60000	6	384
butyl formate	7	20	60000	5	320
ethyl acetate	6	35	51000	8	320
propyl acetate	7	30	60000	8	320
butyl acetate	7	30	63000	8	288
pentyl acetate	8	25	80000	5	320

Table G.2.: Number of CB steps for different dimensionless densities

ρ^*	CB steps per bead
$< 1/2.7$	1
$\geq 1/2.7 \wedge < 1/1.9$	2
$\geq 1/1.9 \wedge < 1/1.6$	3
$\geq 1/1.6 \wedge < 1/1.3$	4
$\geq 1/1.3 \wedge < 1/1.16$	6
$\geq 1/1.16$	8

Configurational Bias Monte Carlo

As mentioned in the main article, the sampling of insertion, deletion, or regrowth steps was enhanced using the configurational biasing (CB) method. Since for the lowest densities the acceptance probability is reasonably high, we perform only one CB step for densities up to a dimensionless density of $\rho^* \leq 1/2.7$, with $\rho^* = \rho/\rho_{\text{max}}$, where ρ_{max} again is the highest density for which we perform simulations. Table G.2 shows the dimensionless densities and the according CB steps.

G.3. Results for formates

Table G.3.: Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization for formates

ethylformate		1 run		
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{\text{lv}}h$ /kJ/kg
320	891	2.18	0.744	412
360	833	7.26	2.73	376
400	770	18.9	7.37	336
440	695	42.8	16.4	284
480	595	94.2	31.7	208
490	562	119	36.9	183
512	336	336	50.9	0
propylformate		1 run		
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{\text{lv}}h$ /kJ/kg
320	864	1.02	0.307	383
360	818	3.8	1.24	358
400	764	10.7	3.68	326
440	705	25.1	8.75	288
480	631	53.7	17.8	237
515	540	107	30.4	171
540	318	318	43	0
butylformate		1 run		
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{\text{lv}}h$ /kJ/kg
320	858	0.435	0.113	372
360	814	1.85	0.533	348
400	768	5.71	1.76	323
440	719	14.2	4.54	293
480	659	31.1	9.86	255
520	586	63.9	18.9	205
545	522	105	27	158
571	307	307	38.2	0

Figure G.1.: n-butyl-formate, n-propyl-formate, ethyl-formate: Comparison of calculated phase equilibrium data from the TAMie force field to quasi-experimental data[137] in a T - ρ -projection.

Table G.4.: Critical properties of simulated formates

substance	simulation results			experimental data [137, 187, 188]		
	T_C /K	p_C /bar	ρ_C /kg/m ³	T_C /K	p_C /bar	ρ_C /kg/m ³
ethyl formate	512	50.9	336	508.4	47.4	323.2
propyl formate	540	43.0	318	538.0	40.6	309.3
butyl formate	571	38.2	307	559.0	35.1	304.0

G.4. Results for acetates

Table G.5.: Tabulated values of liquid densities, vapor pressures and enthalpies of vaporization for acetates

ethyl acetate				
1 run				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
320	863	1.04	0.313	382
360	813	3.88	1.26	355
400	755	11.2	3.74	321
440	690	27.1	9	278
480	607	61.3	18.6	218
505	528	107	27.7	161
528	312	312	38.7	0
propyl acetate				
1 run				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
320	851	0.518	0.134	369
360	805	2.23	0.632	343
400	755	6.91	2.08	314
440	699	17.6	5.36	279
480	632	39.5	11.6	234
520	536	88	22.3	167
525	519	98.7	24.1	156
550	302	302	34.4	0
butylacetate				
1 run				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
320	845	0.222	0.0505	360
360	808	1.12	0.283	339
400	761	3.88	1.05	312
440	711	10.5	2.96	283
480	656	24.3	6.84	247
520	584	51.9	13.7	200
550	509	93.7	21.6	150
578	295	295	31.9	0
pentylacetate				
1 run				
T /K	ρ^l /kg/m ³	ρ^v /kg/m ³	p^{sat} /bar	$\Delta^{lv}h$ /kJ/kg
320	836	0.09	0.0183	355
360	806	0.542	0.123	336
400	767	2.13	0.525	313
440	722	6.23	1.63	287
480	674	15.2	4.04	258
520	616	33	8.59	221
560	541	68.3	16.2	171
575	499	92.2	20.1	143
603	285	285	29.3	0

Table G.6.: Critical properties of simulated acetates

substance	simulation results			experimental data 137 , 139 , 186 , 188–190		
	T_C /K	p_C /bar	ρ_C /kg/m ³	T_C /K	p_C /bar	ρ_C /kg/m ³
ethyl acetate	528	38.7	312	523.3	38.5	308
propyl acetate	550	34.4	302	549.4	33.6	296
butyl acetate	578	31.9	295	576	30.9	–
pentyl acetate	603	29.3	285	–	–	–

G. Supporting information for chapter 5

Figure G.2.: ethyl-acetate, n-propyl-acetate, n-butyl-acetate and n-pentyl-acetate: Comparison of calculated vapor pressures from the TAMie force field to quasi-experimental data[137].

Bibliography

- [1] A. Hemmen and J. Gross. Transferable anisotropic united-atom force field based on the mie potential for phase equilibrium calculations: n-alkanes and n-olefins. *J. Phys. Chem. B*, 119(35):11695–11707, 2015.
- [2] H. Flyvbjerg and H. Petersen. Error estimates on averages of correlated data. *J. Chem. Phys.*, 91:461, 1989.
- [3] D. Weidler and J. Gross. Transferable anisotropic united-atom force field based on the mie potential for phase equilibria: aldehydes, ketones and small cyclic alkanes. *Ind. Eng. Chem. Res.*, 55(46):12123–12132, 2016.
- [4] D. Weidler and J. Gross. Individualized force fields for alkanes, olefins, ethers and ketones based on the transferable anisotropic mie potential. *Fluid Phase Equilib.*, 470:102–108, 2018.
- [5] D. Weidler and J. Gross. Phase equilibria of binary mixtures with alkanes, ketones, and esters based on the transferable anisotropic mie force field. *Fluid Phase Equilib.*, 490:123–132, 2019.
- [6] A. Hemmen, A. Z. Panagiotopoulos, and J. Gross. Grand canonical monte carlo simulations guided by an analytic equation of state – transferable anisotropic mie potentials for ethers. *J. Phys. Chem. B*, 119(23):7087–7099, 2015.
- [7] A. Hemmen. *Neues transferierbares Kraftfeld (TAMie) zur Berechnung von Phasengleichgewichten mit Monte Carlo Simulationen im großkanonischen Ensemble*. PhD thesis, University of Stuttgart, 2018.
- [8] M. Shell and A. Panagiotopoulos. *Methods for examining phase equilibria*, pages 353–387. Springer Series in Chemical Physics. Dec. 2007. ISBN 9783540736172.
- [9] A. Z. Panagiotopoulos. Force field development for simulations of condensed phases.
- [10] M. Müller and J. de Pablo. *Simulation Techniques for Calculating Free Energies*, pages 67–126. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-35273-0.

BIBLIOGRAPHY

- [11] P. Ungerer, C. Nieto-Draghi, B. Rousseau, G. Ahunbay, and V. Lachet[?]. Molecular simulation of the thermophysical properties of fluids: From understanding toward quantitative predictions. *J. Mol. Liq.*, 134(1):71 – 89, 2007. ISSN 0167-7322.
- [12] K. Binder. Computer simulations of critical phenomena and phase behaviour of fluids. *Mol. Phys.*, 108(14):1797–1815, 2010.
- [13] A. Z. Panagiotopoulos. Monte carlo methods for phase equilibria of fluids. *J. Phys.: Condens. Matter*, 12:R25 – R52, 2000.
- [14] R. Sadus. *Molecular Simulations of Fluids: Theory, Algorithms and Object Orientation*. Elsevier Science, 1 edition, 2002.
- [15] M. E. Tuckerman. *Statistical Mechanics: Theory and Molecular Simulation*. Oxford University Press, 2010.
- [16] D. Frenkel and B. Smit. *Understanding Molecular Simulation*. Academic Press, 1996.
- [17] M. Allen and D. Tildesley. *Statistical Mechanics: Theory and Molecular Simulation*. Oxford Science Publications, 1987.
- [18] A. Leach. *Molecular Modelling : Principles and Applications*. Prentice Hall, 2 edition, 2001.
- [19] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 117(19):5179–5197, 1995.
- [20] MultiMedia LLC. The amber project - the amber force fields, 2020.
- [21] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, and C. Simmerling. ff14sb: Improving the accuracy of protein side chain and backbone parameters from ff99sb. *Journal of Chemical Theory and Computation*, 11(8):3696–3713, 2015. PMID: 26574453.
- [22] K. T. Debiec, D. S. Cerutti, L. R. Baker, A. M. Gronenborn, D. A. Case, and L. T. Chong. Further along the road less traveled: Amber ff15ipq, an original protein force field built on a self-consistent physical model. *Journal of Chemical Theory and Computation*, 12(8):3926–3947, 2016. PMID: 27399642.
- [23] L.-P. Wang, K. A. McKiernan, J. Gomes, K. A. Beauchamp, T. Head-Gordon, J. E. Rice, W. C. Swope, T. J. Martínez, and V. S. Pande. Building a more predictive

BIBLIOGRAPHY

- protein force field: A systematic and reproducible route to amber-fb15. *J. Phys. Chem. B*, 121(16):4023–4039, 2017. PMID: 28306259.
- [24] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, and A. D. MacKerell Jr. Charmm36m: an improved force field for folded and intrinsically disordered proteins. *Nat. methods*, 14:71–73, 2017.
- [25] R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, and A. D. MacKerell. Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone [U+03D5], and side-chain 1 and 2 dihedral angles. *J. Chem. Theo. Comp.*, 8(9):3257–3273, 2012.
- [26] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102(18):3586–3616, 1998.
- [27] A. D. Mackerell Jr., M. Feig, and C. L. Brooks III. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comp. Chem.*, 25(11):1400–1415, 2004.
- [28] K. Hart, N. Foloppe, C. M. Baker, E. J. Denning, L. Nilsson, and A. D. MacKerell. Optimization of the charmm additive force field for dna: Improved treatment of the bi/bii conformational equilibrium. *J. Chem. Theo. Comput.*, 8(1):348–362, 2012.
- [29] E. J. Denning, U. D. Priyakumar, L. Nilsson, and A. D. Mackerell Jr. Impact of 2'-hydroxyl sampling on the conformational properties of rna: Update of the charmm all-atom additive force field for rna. *J. Comp. Chem.*, 32(9):1929–1943.
- [30] N. Foloppe and A. D. MacKerell Jr. All-atom empirical force field for nucleic acids: I. parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comp. Chem.*, 21(2):86–104, 2000.
- [31] A. D. MacKerell Jr. and N. K. Banavali. All-atom empirical force field for nucleic acids: Ii. application to molecular dynamics simulations of dna and rna in solution. *J. Comp. Chem.*, 21(2):105–120, 2000.

BIBLIOGRAPHY

- [32] J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, and R. W. Pastor. Update of the charmm all-atom additive force field for lipids: Validation on six lipid types. *J. Phys. Chem. B*, 114(23):7830–7843, 2010.
- [33] M. Schlenkrich, J. Brickmann, A. MacKerell Jr., and M. Karplus. An empirical potential energy function for phospholipids: Criteria for parameter optimization and applications. In *Biological Membranes: A Molecular Perspective from Computation and Experiment*. Birkhäuser, 1996.
- [34] S. E. Feller, D. Yin, R. W. Pastor, and A. D. MacKerell Jr. Molecular dynamics simulation of unsaturated lipids at low hydration: Parametrization and comparison with diffraction studies. *Biophys. J.*, 73:2269–2279, 1997.
- [35] O. Guvench, S. S. Mallajosyula, E. P. Raman, E. Hatcher, K. Vanommeslaeghe, T. J. Foster, F. W. Jamison, and A. D. MacKerell. Charmm additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J. Chem. Theo. Comp.*, 7(10):3162–3180, 2011.
- [36] S. S. Mallajosyula, O. Guvench, E. Hatcher, and A. D. MacKerell. Charmm additive all-atom force field for phosphate and sulfate linked to carbohydrates. *J. Chem. Theo. Comp.*, 8(2):759–776, 2012.
- [37] E. P. Raman, O. Guvench, and A. D. MacKerell. Charmm additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses. *J. Phys. Chem. B*, 114(40):12981–12994, 2010.
- [38] O. Guvench, E. Hatcher, R. M. Venable, R. W. Pastor, and A. D. MacKerell. Charmm additive all-atom force field for glycosidic linkages between hexopyranoses. *J. Chem. Theo. Comp.*, 5(9):2353–2370, 2009.
- [39] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, 118(45):11225–11236, 1996.
- [40] B. Chen and J. I. Siepmann. Transferable potentials for phase equilibria. 3. explicit-hydrogen description of normal alkanes. *J. Phys. Chem. B*, 103(25):5370–5379, 1999.
- [41] N. Rai and J. I. Siepmann. Transferable potentials for phase equilibria. 9. explicit hydrogen description of benzene and five-membered and six-membered heterocyclic aromatic compounds. *J. Phys. Chem. B*, 111(36):10790–10799, 2007.

BIBLIOGRAPHY

- [42] N. Rai and J. I. Siepmann. Transferable potentials for phase equilibria. 10. explicit-hydrogen description of substituted benzenes and polycyclic aromatic compounds. *J. Phys. Chem. B*, 117(1):273–288, 2013.
- [43] E. O. Fetisov and J. I. Siepmann. Prediction of vapor–liquid coexistence properties and critical points of polychlorinated biphenyls from monte carlo simulations with the trapeze force field. *J. Chem. Eng. Data*, 59(10):3301–3306, 2014.
- [44] W. L. Jorgensen, J. D. Madura, and C. J. Swenson. Optimized intermolecular potential functions for liquid hydrocarbons. *J. Am. Chem. Soc.*, 106:6638–6646, 1984.
- [45] W. L. Jorgensen and C. J. Swenson. Optimized intermolecular potential functions for amides and peptides. structure and properties of liquid amides. *J. Am. Chem. Soc.*, 107:569–578, 1985.
- [46] W. L. Jorgensen and C. J. Swenson. Optimized intermolecular potential functions for amides and peptides. hydration of amides. *J. Am. Chem. Soc.*, 107:1489–1496, 1985.
- [47] W. L. Jorgensen. Optimized intermolecular potential functions for liquid alcohols. *J. Phys. Chem.*, 90:1276–1284, 1986.
- [48] W. L. Jorgensen and J. Tirado-Rives. The opls force field for proteins. energy minimization for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, 110(8):1657–1666, 1988.
- [49] J. M. Briggs, T. Matsui, and W. L. Jorgensen. Monte carlo simulations of liquid alkyl ethers with the opls potential functions. *J. Comp. Chem*, 11(8):958–971, 1990.
- [50] M. G. Martin and J. I. Siepmann. Transferable potentials for phase equilibria. 1. united-atom description of n-alkanes. *J. Phys. Chem. B*, 102(14):2569–2577, 1998.
- [51] M. G. Martin and J. I. Siepmann. Novel configurational-bias monte carlo method for branched molecules. transferable potentials for phase equilibria. 2. united-atom description of branched alkanes. *J. Phys. Chem. B*, 103(21):4508–4517, 1999.
- [52] B. Chen, J. J. Potoff, and J. I. Siepmann. Monte carlo calculations for alcohols and their mixtures with alkanes. transferable potentials for phase equilibria. 5. united-atom description of primary, secondary, and tertiary alcohols. *J. Phys. Chem. B*, 105(15):3093–3104, 2001.

BIBLIOGRAPHY

- [53] J. M. Stubbs, J. J. Potoff, and J. I. Siepmann. Transferable potentials for phase equilibria. 6. united-atom description for ethers, glycols, ketones, and aldehydes. *J. Phys. Chem. B*, 108(45):17596–17605, 2004.
- [54] J. I. Siepmann. The transferable potentials for phase equilibria - webpage. Accessed: 2016-03-15.
- [55] E. A. Mueller and G. Jackson. Force-field parameters from the soft-gamma equation of state for use in coarse-grained molecular simulations. *Annu. Rev. Chem. Biomol. Eng.*, 5:405–427, 2014.
- [56] T. Lafitte, A. Apostolakou, C. Avendaño, A. Galindo, C. S. Adjiman, E. A. Müller, and G. Jackson. Accurate statistical associating fluid theory for chain molecules formed from mie segments. *J. Chem. Phys.*, 139(15), 2013.
- [57] C. Avendaño, T. Lafitte, C. S. Adjiman, A. Galindo, E. A. Müller, and G. Jackson. Soft- γ force field for the simulation of molecular fluids: 2. coarse-grained models of greenhouse gases, refrigerants, and long alkanes. *J. Phys. Chem. B*, 117(9):2717–2733, 2013.
- [58] C. Waibel, R. Stierle, and J. Gross. Transferability of cross-interaction parameters for molecular simulation: Vapor-liquid phase equilibria of n-alkane-nitrogen mixtures using the TAMie force field. *Fluid Phase Equilib.*, 456, 2017.
- [59] J. Gross and G. Sadowski. Application of perturbation theory to a hard-chain reference fluid: an equation of state for square-well chains. *Fluid Phase Equilib.*, 168(2):183–199, 2000.
- [60] J. Gross and G. Sadowski. Perturbed-chain saft: An equation of state based on a perturbation theory for chain molecules. *Ind. Eng. Chem. Res.*, 40(4):1244–1260, 2001.
- [61] J. Gross and G. Sadowski*. Application of the perturbed-chain saft equation of state to associating systems. *Ind. Eng. Chem. Res.*, 41(22):5510–5515, 2002.
- [62] J. Gross. An equation-of-state contribution for polar components: Quadrupolar molecules. *AIChE J.*, 51(9):2556–2568.
- [63] D. A. McQuarrie and J. D. Simon. *Molecular Thermodynamics*. University Science Books, 1999. ISBN 978-1-891389-05-4.
- [64] S. Alavi. *Molecular Simulations: Fundamentals and Practice*. Wiley-VCH, 2020. ISBN 978-3-527-34105-4.

BIBLIOGRAPHY

- [65] E. Neyts and A. Bogaerts. Combining molecular dynamics with monte carlo simulations: implementations and applications. *Theor. Chem. Acc.*, 132:1320, 2013.
- [66] G. Raabe. *Molecular Simulation Studies on Thermophysical Properties: With application to Working Fluids*. Springer, 2017. ISBN 978-981-10-3544-9.
- [67] R. Lustig. Microcanonical monte carlo simulation of thermodynamic properties. *J. Chem. Phys.*, 109(20):8816–8828, 1998.
- [68] K. Meier and S. Kabelac. Pressure derivatives in the classical molecular-dynamics ensemble. *J. Chem. Phys.*, 124(6):064104–1–064104–10, 2005.
- [69] A. Z. Panagiotopoulos. Direct determination of phase coexistence properties of fluids by monte carlo simulation in a new ensemble. *Mol. Phys.*, 61(4):813–826, 1987.
- [70] G. Guevara-Carrion, H. Hasse, and J. Vrabec. *Thermodynamic Properties for Applications in Chemical Industry via Classical Force Fields*, pages 201–249. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012. ISBN 978-3-642-24968-6.
- [71] J. J. de Pablo, Q. Yan, and F. A. Escobedo. Simulation of phase transitions in fluids. *Annu. Rev. Phys. Chem.*, 50(1):377–411, 1999.
- [72] A. M. Ferrenberg and R. H. Swendsen. New monte carlo technique for studying phase transitions. *Phys. Rev. Lett.*, 61:2635–2638, Dec 1988.
- [73] A. M. Ferrenberg and R. H. Swendsen. Optimized monte carlo data analysis. *Phys. Rev. Lett.*, 63(12):1195–1198, 1989.
- [74] N. B. Wilding. Critical-point and coexistence-curve properties of the lennard-jones fluid: A finite-size scaling study. *Phys. Rev. E*, 52:602–611, Jul 1995.
- [75] A. Z. Panagiotopoulos, V. Wong, and M. A. Floriano. Phase equilibria of lattice polymers from histogram reweighting monte carlo simulations. *Macromolec.*, 31(3): 912–918, 1998.
- [76] M. Fitzgerald, R. R. Picard, and R. N. Silver. Canonical transition probabilities for adaptive metropolis simulation. *EPL*, 46(3):282, 1999.
- [77] M. Fitzgerald, R. R. Picard, and R. N. Silver. Monte carlo transition dynamics and variance reduction. *J. Stat. Phys.*, 98(1):321–345, 2000.
- [78] J. R. Errington. Direct calculation of liquid-vapor phase equilibria from transition matrix monte carlo simulation. *J. Chem. Phys.*, 118:9915–9925, 2003.

BIBLIOGRAPHY

- [79] V. K. Shen and J. R. Errington. Determination of fluid-phase behavior using transition-matrix monte carlo: Binary lennard-jones mixtures. *J. Chem. Phys.*, 122(6), 2005.
- [80] J. J. Potoff and A. Z. Panagiotopoulos. Critical point and phase behavior of the pure fluid and a lennard-jones mixture. 109(24), 12 1998.
- [81] J. Potoff, J. Errington, and A. Panagiotopoulos. Molecular simulation of phase equilibria for mixtures of polar and non-polar components. *Molecular Physics*, 97(10):1073–1083, 1999.
- [82] Y. Liu, A. Z. Panagiotopoulos, and P. G. Debenedetti. Monte carlo simulations of high-pressure phase equilibria of co₂-h₂o mixtures. *The Journal of Physical Chemistry B*, 115(20):6629–6635, 2011.
- [83] J. I. Siepmann and D. Frenkel. Configurational bias monte carlo: a new sampling scheme for flexible chains. *Mol. Phys.*, 75(1):59–70, 1992.
- [84] B. S. J. S. T.J.H. Vlugt, M.G. Martin and R. Krishna. Improving the efficiency of the configurational-bias monte carlo algorithm. *Mol. Phys.*, 94:727–733, 1998.
- [85] M. N. Rosenbluth and A. W. Rosenbluth. Monte carlo calculation of the average extension of molecular chains. *J.Chem. Phys*, 23(2):356–359, 1955.
- [86] T. J. H. Vlugt, R. Krishna, , and B. Smit*. Molecular simulations of adsorption isotherms for linear and branched alkanes and their mixtures in silicalite. *J. Phys. Chem. B*, 103(7):1102–1118, 1999.
- [87] J. K. Shah and E. J. Maginn. A general and efficient monte carlo method for sampling intramolecular degrees of freedom of branched and cyclic molecules. *J. Chem. Phys.*, 135(13):134121, 2011.
- [88] J. R. Errington and A. Z. Panagiotopoulos. New intermolecular potential models for benzene and cyclohexane. *J. Chem. Phys.*, 111(21):9731–9738, 1999.
- [89] M. D. Macedonia and E. J. Maginn. A biased grand canonical monte carlo method for simulating adsorption using all-atom and branched united atom models. *Mol. Phys.*, 96(9):1375–1390, 1999.
- [90] C. D. Wick, M. G. Martin, and J. I. Siepmann. Transferable potentials for phase equilibria. 4. united-atom description of linear and branched alkenes and alkylbenzenes. *J. Phys. Chem. B*, 104(33):8008–8016, 2000.

BIBLIOGRAPHY

- [91] P. H. Verdier and W. H. Stockmayer. Monte carlo calculations on the dynamics of polymers in dilute solution. *J. Chem. Phys.*, 36(1):227–235, 1962.
- [92] A. Baumgärtner and K. Binder. Monte carlo studies on the freely jointed polymer chain with excluded volume interaction. *J. Chem. Phys.*, 71(6):2541–2545, 1979.
- [93] M. S. Wertheim. Fluids with highly directional attractive forces. i. statistical thermodynamics. *J. Stat. Phys.*, 35(1):19–34, Apr 1984. ISSN 1572-9613.
- [94] M. S. Wertheim. Fluids with highly directional attractive forces. ii. thermodynamic perturbation theory and integral equations. *J. Stat. Phys.*, 35(1):35–47, Apr 1984. ISSN 1572-9613.
- [95] M. S. Wertheim. Fluids with highly directional attractive forces. iii. multiple attraction sites. *J. Stat. Phys.*, 42(3):459–476, Feb 1986. ISSN 1572-9613.
- [96] M. S. Wertheim. Fluids with highly directional attractive forces. iv. equilibrium polymerization. *J. Stat. Phys.*, 42(3):477–492, Feb 1986. ISSN 1572-9613.
- [97] W. G. Chapman, G. Jackson, and K. E. Gubbins. Phase equilibria of associating fluids. *Mol. Phys.*, 65(5):1057–1079, 1988.
- [98] W. G. Chapman, K. E. Gubbins, G. Jackson, and M. Radosz. New reference equation of state for associating liquids. *Ind. Eng. Chem. Res.*, 29(8):1709–1721, 1990.
- [99] A. Gil-Villegas, A. Galindo, P. J. Whitehead, S. J. Mills, G. Jackson, and A. N. Burgess. Statistical associating fluid theory for chain molecules with attractive potentials of variable range. *J. Chem. Phys.*, 106(10):4168–4186, 1997.
- [100] C. Avendaño, T. Lafitte, A. Galindo, C. S. Adjiman, G. Jackson, and E. A. Müller. Saft- γ force field for the simulation of molecular fluids. 1. a single-site coarse grained model of carbon dioxide. *J. Phys. Chem. B*, 115(38):11154–11169, 2011. PMID: 21815624.
- [101] Y. Tang. A saft model for associating lennard-jones chain mixtures. *Mol. Phys.*, 100(7):1033–1047, 2002.
- [102] J. A. Barker and D. Henderson. Perturbation theory and equation of state for fluids: The square-well potential. *J. Chem. Phys.*, 47(8):2856–2861, 1967.
- [103] J. A. Barker and D. Henderson. Perturbation theory and equation of state for fluids. ii. a successful theory of liquids. *J. Chem. Phys.*, 47(11):4714–4721, 1967.
- [104] J. Gross and J. Vrabec. An equation-of-state contribution for polar components: Dipolar molecules. *AIChE J.*, 52(3):1194–1204.

BIBLIOGRAPHY

- [105] M. Kleiner and J. Gross. An equation of state contribution for polar components: Polarizable dipoles. *AIChE J.*, 52(5):1951–1961, 2006.
- [106] M. R. Shirts and J. D. Chodera. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.*, 129(12):124105, 2008.
- [107] T. van Westen, T. J. H. Vlugt, and J. Gross. Determining force field parameters using a physically based equation of state. *J. Phys. Chem. B*, 115(24):7872–7880, 2011.
- [108] C. S. Schacht, T. J. H. Vlugt, and J. Gross. Using an analytic equation of state to obtain quantitative solubilities of CO_2 by molecular simulation. *J. Phys. Chem. Lett.*, 2(5):393–396, 2011.
- [109] D. Berthelot. Sur le mélange des gaz. *C. R. Hebd. Seanc. Acad. Sci. (Paris)*, 126: 1703, 1898.
- [110] H. A. Lorentz. Ueber die anwendung des satzes vom virial in der kinetischen theorie der gase. *Ann. Phys.*, 248:127, 1881.
- [111] J. J. Potoff and D. A. Bernard-Brunel. Mie potentials for phase equilibria calculations: Application to alkanes and perfluoralkanes. *J. Phys. Chem. B*, 113: 14725–14731, 2009.
- [112] C. D. Wick, J. M. Stubbs, N. Rai, and W. J. Ilya Siepmann. Transferable potentials for phase equilibria. 7. primary, secondary, and tertiary amines, nitroalkanes and nitrobenzene, nitriles, amides, pyridine, and pyrimidine. *J. Phys. Chem. B*, 109 (40):18974–18982, 2005.
- [113] S. J. Keasler, S. M. Charan, C. D. Wick, I. G. Economou, and J. I. Siepmann. Transferable potentials for phase equilibria-united atom description of five- and six-membered cyclic alkanes and ethers. *J. Phys. Chem. B*, 116(36):11234–11246, 2012.
- [114] J. R. Errington and A. Z. Panagiotopoulos. New intermolecular potential models for benzene and cyclohexane. *J. Chem. Phys.*, 111(21):9731–9738, 1999.
- [115] S. K. Nath, F. A. Escobedo, and J. J. de Pablo. On the simulation of vapor-liquid equilibria for alkanes. *J. Chem. Phys.*, 108(23):9905–9911, 1998.
- [116] R. Khare, A. K. Sum, S. K. Nath, and J. J. de Pablo. Simulation of vapor-liquid phase equilibria of primary alcohols and alcohol-alkane mixtures. *J. Phys. Chem. B*, 108(28):10071–10076, 2004.

BIBLIOGRAPHY

- [117] S. Toxvaerd. Molecular dynamics calculation of the equation of state of alkanes. *J. Chem. Phys.*, 93(6):4290–4295, 1990.
- [118] B. Smit, S. Karaborni, and J. I. Siepmann. Computer simulations of vapor-liquid phase equilibria on n-alkanes. *J. Chem. Phys.*, 102(5):2126–2140, 1995.
- [119] P. Ungerer, C. Beauvais, J. Delhommelle, A. Boutin, and B. Rousseau. Optimization of the anisotropic united atoms intermolecular potential for n-alkanes, gibbs ensemble. *J. Chem. Phys.*, 112:5499–5510, 2000.
- [120] E. Bourasseau, P. Ungerer, and A. Boutin. Prediction of equilibrium properties of cyclic alkanes by monte carlo simulations - new anisotropic united atoms intermolecular potentials - new transfer bias method. *J. Phys. Chem. B*, 106(21):5483–5491, 2002.
- [121] E. Bourasseau, M. Haboudou, A. Boutin, A. H. Fuchs, and P. Ungerer. New optimization method for intermolecular potentials: Optimization of a new anisotropic united atoms potential for olefins: Prediction of equilibrium properties. *J. Chem. Phys.*, 118(7):3020–3034, 2003.
- [122] R. O. Contreras-Camacho, P. Ungerer, A. Boutin, and A. D. Mackie. Optimized intermolecular potential for aromatic hydrocarbons based on anisotropic united atoms. 1. benzene. *J. Phys. Chem. B*, 108(37):14109–14114, 2004.
- [123] C. Nieto-Draghi, P. Bonnaud, and P. Ungerer. Anisotropic united atom model including the electrostatic interactions of methylbenzenes. i. thermodynamic and structural properties. *J. Phys. Chem. C*, 111(43):15686–15699, 2007.
- [124] J. Pérez-Pellitero, P. Ungerer, and A. D. Mackie. An anisotropic united atoms (aua) potential for thiophenes. *J. Phys. Chem. B*, 111(17):4460–4466, 2007.
- [125] N. Ferrando, V. Lachet, J.-M. Teuler, and A. Boutin. Transferable force field for alcohols and polyalcohols. *J. Phys. Chem. B*, 113(17):5985–5995, 2009.
- [126] N. Ferrando, V. Lachet, and A. Boutin. Monte carlo simulations of mixtures involving ketones and aldehydes by a direct bubble pressure calculation. *J. Phys. Chem. B*, 114(26):8680–8688, 2010.
- [127] N. Ferrando, V. Lachet, J. Pérez-Pellitero, A. D. Mackie, P. Malfreyt, and A. Boutin. A transferable force field to predict phase equilibria and surface tension of ethers and glycol ethers. *J. Phys. Chem. B*, 115(36):10654–10664, 2011.

BIBLIOGRAPHY

- [128] N. Ferrando, V. Lachet, and A. Boutin. Transferable force field for carboxylate esters: Application to fatty acid methylic ester phase equilibria prediction. *J. Phys. Chem. B*, 116(10):3239–3248, 2012.
- [129] G. A. Orozco, C. Nieto-Draghi, A. D. Mackie, and V. Lachet. Transferable force field for equilibrium and transport properties in linear, branched, and bifunctional amines i. primary amines. *J. Phys. Chem. B*, 115(49):14617–14625, 2011.
- [130] G. A. Orozco, C. Nieto-Draghi, A. D. Mackie, and V. Lachet. Transferable force field for equilibrium and transport properties in linear and branched monofunctional and multifunctional amines. ii. secondary and tertiary amines. *J. Phys. Chem. B*, 116(21):6193–6202, 2012.
- [131] G. A. Orozco, V. Lachet, C. Nieto-Draghi, and A. D. Mackie. A transferable force field for primary, secondary, and tertiary alkanolamines. *J. Chem. Theory Comput.*, 9(4):2097–2103, 2013.
- [132] S. Kranias, D. Pattou, B. Levy, and A. Boutin. An optimized potential for phase equilibria calculation for ketone and aldehyde molecular fluids. *Phys. Chem. Chem. Phys.*, 5:4175–4179, 2003.
- [133] J. J. Potoff and G. Kamath. Mie potentials for phase equilibria: Application to alkenes. *J. Chem. Eng. Data*, 59(10):3144–3150, 2014.
- [134] T. Lafitte, C. Avendaño, V. Papaioannou, A. Galindo, C. S. Adjiman, G. Jackson, and E. A. Müller. Saft- γ force field for the simulation of molecular fluids: 3. coarse-grained models of benzene and hetero-group models of n-decylbenzene. *Mol. Phys.*, 110(11-12):1189–1203, 2012.
- [135] T. Stiegler and R. J. Sadus. Molecular simulation of fluids with non-identical intermolecular potentials: Thermodynamic properties of 10-5 + 12-6 mie potential binary mixtures. *J. Chem. Phys.*, 142(8):084504, 2015.
- [136] P. van der Ploeg and H. J. C. Berendsen. Molecular dynamics simulation of a bilayer membrane. *J. Chem. Phys.*, 76(6):3271–3276, 1982.
- [137] R. L. Rowley, W. V. Wilding, J. L. Oscarson, Y. Yang, N. A. Zundel, T. E. Daubert, and R. P. Danner. Dippr data compilation of pure chemical properties. design institute for physical properties, 2009.
- [138] Korean thermophysical properties data bank, 2016.
- [139] Dortmund data bank, 2015.

BIBLIOGRAPHY

- [140] J.-S. Wang, T. K. Tay, and R. H. Swendsen. Transition matrix monte carlo reweighting and dynamics. *Phys. Rev. Lett.*, 82:476–479, Jan 1999.
- [141] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov. New approach to monte carlo calculation of the free energy: Method of expanded ensembles. *J. Chem. Phys.*, 96(3):1776–1783, 1992.
- [142] P. Virnau and M. Müller. Calculation of free energy through successive umbrella sampling. *J. Chem. Phys.*, 120(23):10925–10930, 2004.
- [143] S. J. Keasler, S. M. Charan, C. D. Wick, I. G. Economou, and J. I. Siepmann. Transferable potentials for phase equilibria – united atom description of five- and six-membered cyclic alkanes and ethers. *J. Phys. Chem. B*, 116(36):11234–11246, 2012.
- [144] Y. M. Muñoz-Muñoz, G. Guevara-Carrion, M. Llano-Restrepo, and J. Vrabec. Lennard-jones force field parameters for cyclic alkanes from cyclopropane to cyclohexane. *Fluid Phase Equilib.*, 404:150 – 160, 2015. ISSN 0378-3812.
- [145] B. Neubauer, A. Boutin, B. Tavitian, and A. H. Fuchs. Gibbs ensemble simulations of vapour-liquid phase equilibria of cyclic alkanes. *Mol. Phys.*, 97(6):769–776, 1999.
- [146] G. Kamath, G. Georgiev, , and J. J. Potoff. Molecular modeling of phase behavior and microstructure of acetone-chloroform-methanol binary mixtures. *J. Phys. Chem. B*, 109(41):19463–19473, 2005.
- [147] S. Toxvaerd. Equation of state of alkanes ii. *J. Chem. Phys.*, 107(13):5197–5204, 1997.
- [148] J. Pérez-Pellitero, E. Bourasseau, I. Demachy, J. Ridard, P. Ungerer, and A. D. Mackie. Anisotropic united-atoms (aua) potential for alcohols. *J. Phys. Chem. B*, 112(32):9853–9863, 2008.
- [149] J. R. Mick, M. S. Barhaghi, B. Jackman, K. Rushaidat, L. Schwiebert, and J. J. Potoff. Optimized mie potentials for phase equilibria: Application to noble gases and their mixtures with n-alkanes. *J. Chem. Phys.*, 143(11):114504, 2015.
- [150] M. S. Barhaghi, J. R. Mick, and J. J. Potoff. Optimised mie potentials for phase equilibria: application to alkynes. *Mol. Phys.*, 115(9-12):1378–1388, 2017.
- [151] A. F. Ghobadi and J. R. Elliott. Adapting saft- γ perturbation theory to site-based molecular dynamics simulation. i. homogeneous fluids. *J. Chem. Phys.*, 139(23), 2013.

BIBLIOGRAPHY

- [152] A. F. Ghobadi and J. R. Elliott. Adapting soft- γ perturbation theory to site-based molecular dynamics simulation. ii. confined fluids and vapor-liquid interfaces. *J. Chem. Phys.*, 141(2), 2014.
- [153] A. F. Ghobadi and J. R. Elliott. Adapting soft- γ perturbation theory to site-based molecular dynamics simulation. iii. molecules with partial charges at bulk phases, confined geometries and interfaces. *J. Chem. Phys.*, 141(9):094708, September 2014.
- [154] I. D. Gospodinov and F. A. Escobedo. Bridging continuum and statistical thermodynamics via equations of state and the density of states. *J. Chem. Phys.*, 120(22):10699–10710, 2004.
- [155] I. D. Gospodinov and F. A. Escobedo*. Probability density of macrostates and density of states for multi-component mixtures from semi-empirical equations of state. *Mol. Phys.*, 103(21-23):3115–3124, 2005.
- [156] J. M. C. Sanchez, T. Danner, and J. Gross. Grand canonical monte carlo simulations of vapor-liquid equilibria using a bias potential from an analytic equation of state. *J. Chem. Phys.*, 138(23):234106, 2013.
- [157] B. Haas. Simulation verzweigter alkane im großkanonischen ensemble und optimierung der charakteristischen intermolekularen wechselwirkungsparameter. Master's thesis.
- [158] D. Frenkel, G. Mooij, and B. Smit. Novel scheme to study structural and thermal properties of continuously deformable molecules. *J. Phys. Condens. Matter*, 4(12):3053, 1992.
- [159] E. H. T. Boublik, V. Fried. *The vapour pressure of pure substances*. Elsevier, 1973.
- [160] D. Stull. Vapor pressure of pure substances. *Ind. Eng. Chem.*, 39:517–540, 1947.
- [161] Selected values of properties of chemical compounds, thermodynamics research center, texas a&m university, college station, texas.
- [162] E. Washburn. *International Critical Tables of Numerical Data, Physics, Chemistry, and Technology*. McGraw-Hill, 1926–1933. 7 Vols + Index.
- [163] D. Ambrose, J. Ellender, C. Sprake, and R. Townsend. Thermodynamic properties of organic oxygen compounds - xliii. vapour pressures of some ethers. *J. Chem. Thermodyn.*, 8:165, 1976.
- [164] D. O. Hanson and M. van Winkle. Alteration of the relative volatility of n-hexane–1-hexene by oxygenated and chlorinated solvents. *J. Chem. Eng. Data*, 12(3):319–325, 1967.

BIBLIOGRAPHY

- [165] J. M. Stubbs, J. J. Potoff, and J. I. Siepmann. Transferable potentials for phase equilibria. 6. united-atom description for ethers, glycols, ketones, and aldehydes. *J. Phys. Chem. B*, 108(45):17596–17605, 2004.
- [166] S. Eckelsbach, T. Janzen, A. Köster, S. Miroshnichenko, Y. M. Muñoz-Muñoz, and J. Vrabec. Molecular models for cyclic alkanes and ethyl acetate as well as surface tension data from molecular simulation. In W. E. Nagel, D. H. Kröner, and M. M. Resch, editors, *High Performance Computing in Science and Engineering 14*, pages 645–659, Cham, 2015. Springer International Publishing.
- [167] G. Kamath, J. Robinson, and J. J. Potoff. Application of trappe-ua force field for determination of vapor-liquid equilibria of carboxylate esters. *Fluid Phase Equilib.*, 240(1):46 – 55, 2006.
- [168] K. A. Maerzke, N. E. Schultz, R. B. Ross, and J. I. Siepmann. Trappe-ua force field for acrylates and monte carlo simulations for their mixtures with alkanes and alcohols. *J. Phys. Chem. B*, 113(18):6415–6425, 2009.
- [169] Turbomole v6.6 2014, 2014. A development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.
- [170] T. Vlugt, R. Krishna, and B. Smit. Molecular simulations of adsorption isotherms for linear and branched alkanes and their mixtures in silicalite. *J. Phys. Chem. B*, 103(7):1102–1118, 1999.
- [171] J. K. Shah, E. Marin-Rimoldi, R. G. Mullen, B. P. Keene, S. Khan, A. S. Paluch, N. Rai, L. L. Romanielo, Rosch, T. W., Yoo, Brian, Maginn, and E. J. Cassandra: An open source monte carlo package for molecular simulation. *J. Comput. Chem.*, 38(19):1727–1739, 2017.
- [172] G. Shealy and S. Sandler. *J. Chem. Thermodyn.*, 17:143–150, 1985.
- [173] V. Aristovich, A. Morachevskii, and I. Sabylin. *J. Appl. Chem. USSR*, 38(12):2633–2638, 1965.
- [174] W. A. Scheller and S. V. N. Rao. Isothermal vapor-liquid equilibrium data for system heptane-2-pentanone at 90.deg. *J. Chem. Eng. Data*, 18(2):223–225, 1973.
- [175] O. Kiyohara, G. Benson, and J.-P. Grolier. Thermodynamic properties of binary mixtures containing ketones i. excess enthalpies of some aliphatic ketones + n-hexane, + benzene, and + tetrachloromethane. *J. Chem. Thermodyn.*, 9:315–323, 1977.

BIBLIOGRAPHY

- [176] S. Murakami, K. Amaya, and R. Fujishiro. Heats of mixing for binary mixtures. the energy of hydrogen bonding between alcohol and ketone molecules. *Bull. Chem. Soc. Jpn.*, 37(12):1776–1780, 1964.
- [177] Y.-L. Huang, M. Heilig, H. Hasse, and J. Vrabec. Vapor–liquid equilibria of hydrogen chloride, phosgene, benzene, chlorobenzene, ortho-dichlorobenzene, and toluene by molecular simulation. *AIChE J.*, 57(4):1043–1060, 2011.
- [178] H. Hiraoka. Unknown. *Kogyo Kagaku Zasshi*, 62:921–926, 1959.
- [179] J. Fischer and A. Weiss. Vii. viscosity, excess volumes, and self-diffusion of binary mixtures of donor-acceptor and of ccl4-carboxylic acid systems. *Ber.Bunsen-Ges.Phys.Chem.*, 90:1141–1153, 1986.
- [180] A. Collings and R. Mills. Temperature-dependence of self-diffusion for benzene and carbon tetrachloride. *Trans.Faraday Soc.*, 66:2761–2766, 1970.
- [181] H. Ertl and F. Dullien. Self-diffusion and viscosity of some liquids as a function of temperature. *AIChE J.*, 19:1215–1223, 1973.
- [182] E. Sauer, M. Stavrou, and J. Gross. Comparison between a homo- and a heterosegmented group contribution approach based on the perturbed-chain polar statistical associating fluid theory equation of state. *Ind. Eng. Chem. Res.*, 53(38):14854–14864, 2014.
- [183] B. F. Abu-Sharkh, A. Sunaidi, and E. Z. Hamad. Thermodynamic perturbation theory for fused sphere hard chain fluids using nonadditive interactions. *J. Chem. Phys.*, 120(12):5795–5801, March 2004. ISSN 0021-9606.
- [184] B. A. Berg and R. C. Harris. From data to probability densities without histograms. *Comp. Phys. Commun.*, 179:443–448, 2008.
- [185] N. B. Wilding. Simulation studies of fluid critical behaviour. *Journal of Physics: Condensed Matter*, 9(3):585, 1997.
- [186] N. C. WebBook. <http://webbook.nist.gov/chemistry>, may 2013.
- [187] V. Majer and V. Svoboda. *Blackwell Scientific Publications*, page 300, 1985.
- [188] S. Young. *Sci.Proc.Roy.Dublin Soc.*, 12:374–443, 1909-1910.
- [189] T. Hu, Z. Qin, G. Wang, X. Hou, and J. Wang. *J.Chem.Eng.Data*, 49:1809–1814, 2004.
- [190] S. Quadri and A. Kudchadker. *J.Chem.Thermodyn.*, 23:129–134, 1991.