Machine Learning of a Density Functional for Anisotropic Patchy Particles

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Alessandro Rodolfo Simon
aus Werneck

Tübingen
2025

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	23.05.2025
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Martin Oettel
2. Berichterstatter:	Prof. Dr. Roland Roth

List of Publications

The results obtained in this thesis led to the following publications:

Machine Learning of a Density Functional for Anisotropic Patchy Particles
Alessandro Simon, Jens Weimar, Georg Martius and Martin Oettel, *Journal of Chemical Theory and Computation 20.3 (2024), pp. 1062–1077*, doi:10.1021/acs.jctc.3c01238

Contributions: The project was initiated by AS and MO. AS developed the theory, implemented, ran, and analyzed the experiments, with input from MO and GM. JW implemented the data analysis software. AS and MO wrote the first draft with input from GM, and all authors contributed to the final version of the paper.

Machine Learning approaches to classical density functional theory
Alessandro Simon and Martin Oettel, *te Vrugt, M. (Ed.), Artificial Intelligence and Intelligent Matter, to be published by Springer (Cham)*, arXiv:2406.07345

Contributions: The project was initiated by AS and MO. AS wrote the first draft with input from MO, and all authors contributed to the final version of the paper.

The orientational structure of a model patchy particle fluid: simulations, integral equations, density functional theory and machine learning
Alessandro Simon, Luc Belloni and Daniel Borgis, Martin Oettel, *The Journal of Chemical Physics 162.3 (2025), p. 034503*, doi:10.1063/5.0248694

Contributions: The project was initiated by AS and MO. AS ran the MC and ML experiments and analyzed the data. DB performed the MDFT calculations and LB solved the integral equations. AS and MO wrote the first draft and all authors contributed to the final version of the paper.

SRBench++: Principled Benchmarking of Symbolic Regression With Domain-Expert Interpretation
F. O. de Franca, M. Virgolin, M. Kommenda, M. S. Majumder, M. Cranmer, G. Espada, L. Ingelse, A. Fonseca, M. Landajuela, B. Petersen, R. Glatt, N. Mundhenk, C. S. Lee, J. D. Hochhalter, D. L. Randall , P. Kamienny, H. Zhang, G. Dick, **A. Simon**, B. Burlacu, Jaan Kasak, Meera Machado, Casper Wilstrup, W. G. La Cava *IEEE Transactions on Evolutionary Computation*, doi:10.1109/TEVC.2024.3423681

Contributions: FdF, MV, MK, and WLC designed and conducted the competition and analyzed the results. MM served as domain expert for the real-world track. The remaining authors (including AS) participated in the competition and contributed to this manuscript after results were announced.

Contents

Acronyms

RDF	radial distribution function
DCF	direct correlation function
ODF	orientation distribution function
MC	Monte Carlo
OZ	Ornstein–Zernike
FMT	fundamental measure theory
HS	hard sphere
KF	Kern–Frenkel
TPTn	Wertheims thermodynamic perturbation theory of order n
EOS	equation of state
SC	self-consistent

Symbols

$x \sim \mathcal{U}$	uniformly distributed random variable x
$\hat{\mathbf{r}}$	normalized vector on the unit sphere
Ω	shorthand for three Euler angles ϕ, θ, χ
$D^l_{mn}(\phi, \theta, \chi)$	Wigner D-matrix
$d^l_{mn}(\phi, \theta, \chi)$	small Wigner d-matrix
$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$	Wigner $3j$-symbol

Summary

In this work we investigate the homogeneous and inhomogeneous structure of an anisotropic associating liquid, as realized by a Kern–Frenkel potential, through means of computer simulations and density functional theory (DFT). Current state-of-the-art functionals for these systems are based on fundamental measure theory and Wertheim's thermodynamic perturbation theory. They are able to achieve rather accurate results regarding the equation of state and the inhomogeneous structure when restricted to moderately high temperatures and "simple" external potentials.

At low temperatures however, the bonding probability of the particles in the associating liquid increases, leading to a strong orientational ordering of the particles. This results in inaccuracies in the predictions, that can be attributed to the isotropic treatment of the fluid. Motivated by this fact, we propose the explicit treatment of the orientational degrees of freedom of the particles using an expansion of the relevant quantities in Wigner D-matrices. To this end, we examine how the symmetry of the investigated two-body potential influences the expansion coefficients and how the numerical treatment can be simplified by finding invariant basis functions.

After this, the orientation distribution function of the inhomogeneous liquid between two hard walls and around a spherical tracer particle is examined, showing a relatively strong order. In order to treat the system with the DFT formalism we first derive the necessary equations for the case of an anisotropic liquid. Then we apply the popular random phase approximation as a reference for further approaches. We start by proposing a mean-field ansatz incorporating the orientational moments of the density distribution. The necessary interaction potentials are found using a machine-learning based approach, by parametrizing the potentials using differentiable weights. These are then optimized using Monte Carlo simulation data, collected at different state points and a procedure that guarantees the self-consistency of the resulting Euler–Lagrange equations. The final mean-field functional improves on the performance of previous treatments, however numerical stability quickly degrades when stepping outside the training region.

As a second approach, we propose keeping the functional form of the isotropic functional but learning improved weight functions that enter through the weighted densities. While in this case we cannot make predictions about the orientational degrees of freedom of the system, we gain improved numerical stability through keeping the known functional form of the reference functional. Apart from more accurate density profiles, we can also observe that the correlation function resulting from this ansatz and training procedure is much closer to its real value.

Deutsche Zusammenfassung

In dieser Arbeit untersuchen wir die homogene und inhomogene Struktur einer anisotropen assoziierenden Flüssigkeit, wie sie durch ein Kern–Frenkel-Potential realisiert wird. Wir verwenden hierfür Computersimulationen und Dichtefunktionaltheorie (DFT). Die aktuellen Standard-Funktionale für diese Systeme basieren auf der fundamentalen Maßtheorie (fundamental measure theory) und Wertheims thermodynamischer Störungstheorie (thermodynamic perturbation theory). Diese sind in der Lage, recht genaue Ergebnisse hinsichtlich der Zustandsgleichung und der inhomogenen Struktur zu liefern, wenn sie auf moderat hohe Temperaturen und „einfache“ externe Potentiale beschränkt sind.

Bei niedrigen Temperaturen jedoch erhöht sich die Bindungswahrscheinlichkeit der Teilchen in der assoziierenden Flüssigkeit, was zu einer starken orientierungsspezifischen Ordnung der Teilchen führt. Dies resultiert in Ungenauigkeiten bei den Vorhersagen, die auf die isotrope Behandlung des Fluids zurückzuführen sind. Angesichts dieser Tatsache schlagen wir vor, die orientierungsabhängigen Freiheitsgrade der Teilchen explizit zu behandeln, indem relevante Größen in Wigner-D-Matrizen entwickelt werden. Hierzu untersuchen wir, wie die Symmetrie des betrachteten Zweikörperpotentials die Entwicklungskoeffizienten beeinflusst und wie die numerische Behandlung durch das Bestimmen invarianter Basisfunktionen vereinfacht werden kann.

Anschließend wird die Verteilungsfunktion der Orientierung für die inhomogene Flüssigkeit zwischen zwei harten Wänden und um ein sphärisches Tracer-Teilchen untersucht, welche eine relativ starke Ordnung aufzeigt. Um das System mit dem DFT-Formalismus zu behandeln, leiten wir zunächst die notwendigen Gleichungen für den Fall eines anisotropen Fluids her. Danach wenden wir die beliebte Random-Phase-Approximation als Referenzpunkt für weitere Ansätze an.

Wir beginnen mit dem Vorschlag eines Mean-Field Ansatzes, der die orientierungsabhängigen Momente der Dichteverteilung integriert. Die notwendigen Wechselwirkungspotentiale werden mithilfe maschinellen Lernens gefunden, indem diese Potentiale durch differenzierbare Gewichte parametrisiert und anschließend mittels Vergleich zu Monte-Carlo-Simulationen optimiert werden, die bei verschiedenen

Zustandspunkten gesammelt wurden. Ein Verfahren gewährleistet die Selbstkonsistenz der resultierenden Euler–Lagrange-Gleichungen. Das endgültige Mean-Field-Funktional verbessert die Genauigkeit im Vergleich zu vorherigen Ansätzen, jedoch verschlechtert sich die numerische Stabilität schnell, wenn man den Trainingsbereich verlässt.

Als zweiten Ansatz schlagen wir vor, die Funktionsform des isotropen Funktionals beizubehalten, aber verbesserte Gewichtsfunktionen zu lernen, die über die gewichteten Dichten eingehen. Obwohl wir mit diesem Ansatz keine Vorhersagen über die orientierungsabhängigen Freiheitsgrade des Systems treffen können, gewinnen wir an numerischer Stabilität durch die Beibehaltung der bekannten Funktionsform des Referenz-Funktionals. Neben genaueren Dichteprofilen lässt sich beobachten, dass die Korrelationsfunktion aus diesem Ansatz und dem Trainingsverfahren viel näher am tatsächlichen Wert liegt.

1 Introduction

Our aim in the following will be to find an accurate description of a system consisting of a large number of particles. They interact with each other through an anisotropic two-body potential that can serve as a model for more realistic liquids like water. In addition to the interparticle interactions, there is also an external potential that acts on the particles and makes their density distribution become inhomogeneous. Our study lasts to varying degrees on three main pillars. The principal one, computer simulations, allow us to perform idealized *in-silico* experiments. In contrast to experiments performed in the real world, we have virtually unlimited knowledge and control over the system.

Fundamentally however, having a real understanding of the system is a more complicated matter. It consists of quantitative knowledge of the different effects and forces that are present in the system, their relative strength to each other, the range of validity of certain approximations and so own. Due to the large number of degrees of freedom, a direct treatment of the system is rather hopeless, and we need to consider the system in a statistical context. This brings us to the second pillar, density functional theory (DFT) — here in the context of classical systems — which allows us to treat many-body systems in a "coarse-grained" way, in the sense that the observation of individual particles becomes less important, while putting the focus on averaged properties like the number density, in this way reducing the complexity of the treatment.

Finally, there is the theory of integral equations, which allows us to gain insights into the structure of the fluid. Except for some few cases, solutions have to be determined numerically, resting on so-called closure relations, which quantify how certain correlation functions of the liquid behave relative to each other. Depending on the interparticle potential some closures give better results than others, and it is a-priori not clear which one is the best. In the case of anisotropic liquids, the numerical treatment of integral equations is made more difficult by the slowly converging expansion in rotational degrees of freedom.

As already mentioned, our model system is that of an anisotropic associating liquid. This is very different from classical fluid models like the hard-sphere or the Lennard-

Jones model that are isotropic and hence only depend on separation between the particles. Due to the small bonding volume (the phase space volume where two particles are in a bonded state), mean-field or other perturbative approximations fail at describing the behavior observed in simulations. In fact, it turns out that patchy particles, in our case realized by the Kern–Frenkel potential, possess a quite interesting phase diagram and exhibit unusual phenomena like that of an "empty liquid" [Bia+06; HTG11].

The most accurate DFT treatment of patchy particles rests on the so-called thermodynamic perturbation theory (TPT) developed by Wertheim [Wer84a; Wer84b; Wer86a; Wer86b] about four decades ago. The theory produces a perturbative expression for the Helmholtz free energy of a fluid consisting of hard spheres, decorated with sites that interact attractively with the sites on other spheres. The precise geometric details of the interaction do not enter, they are in a sense, "integrated out". Starting from this expression for the free energy different generalizations to the inhomogeneous case were proposed. One of the most accurate is that of the Stopper *et al.* [Sto+18], based on the work by Yu and Wu [YW02a], using the fundamental measure framework, originally developed for hard-sphere systems.

In TPT1 it is assumed that the particles' orientations are distributed isotropically, simplifying the treatment considerably. It is easy to see however, that this approximation will become inaccurate once the temperature of the system is lowered, since this leads to an increase of the bonding probability, making the particles have preferred orientations.

The next logical step to improve the accuracy of the predictions is to include the orientational degrees of freedom. Since the orientational configuration of a (general, non-symmetric) particle is described by three real numbers, this increases the state space by three dimensions, making a numerical treatment on discretized grids too computationally expensive. An alternative to this is the expansion of the relevant degrees of freedom in the so-called Wigner D-matrices, which can be seen as a generalization of spherical harmonics. This expansion reduces the dimensionality of the discretized grids to the number of physical spatial dimension, which in planar or cylindrical symmetry is one. The downside is that the expansion consists of an infinite sum of terms and is potentially slowly converging. An analysis of the relevant expansions coefficients is necessary to reduce the calculations to the minimum.

The initial ansatz for an improved functional will be that of the isotropic reference functional plus a mean-field interaction that takes into account the orientational moments of the density distribution.

Our goal will be to determine the mean-field kernels or interaction potentials by means of machine learning (ML). For this, it is necessary to collect sufficient amounts of training data that is later used to "fit" the model. This is in a sense an *inverse problem* since through the DFT formalism it is rather easy and quick to find predictions of a given model, but we do not know which model, or rather model parameters, give results that are the closest to data coming from computer simulations (the "ground truth").

The inverse problem described here is complicated by the nature of the model prediction, which results from solving a fixed point (FP) equation. Consequently, evaluating the model prediction directly is not feasible; instead, it requires iterative methods to find a solution. This complexity poses additional challenges for optimizing model parameters, typically achieved through gradient descent. In gradient descent optimization, computing the necessary gradient information of the loss function involves back-propagation. However, back-propagation assumes that the network's mapping from input to output consists of a series of differentiable transformations. The iterative nature required by FP equations adds an extra layer of difficulty to this process, making parameter optimization more challenging than in cases where predictions can be directly computed.

Finally, we explore the possibility of using isotropic reference functionals with improved weighted densities. The weights that result from fundamental measure theory are based on the geometry of hard spheres. Meanwhile, the attractive patches that are attached to the hard spheres extend outside the sphere, therefore it is sensible to increase the range of the weights that are used for calculating the weighted densities. Keeping the known functional form of the reference functional can help to improve the numerical stability of the machine-learned functional.

2 Theoretical basis

> If it looks like a tensor, talks like a tensor, and transforms like a tensor, then it probably *is* a tensor.
>
> — Duck test

2.1 Statistical mechanics

We consider a closed system of N particles of mass m, enclosed in a volume V. The particles' positions are encoded in the vector $\mathbf{r}^N$, and the classical momenta $\mathbf{p}^N$. The total energy of the system is given by the Hamiltonian

$$\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) = K_N(\mathbf{p}^N) + U_N(\mathbf{r}^N) + \Phi(\mathbf{r}^N) \tag{2.1}$$

that means, it is split up in the kinetic energy $K_N(\mathbf{p}^N) = \sum_i |\mathbf{p}_i|^2/2m$, the intermolecular potential energy $U_N(\mathbf{r}^N) = \frac{1}{2}\sum_{i<j} u(\mathbf{r}_i, \mathbf{r}_j)$ and the external potential energy $\Phi(\mathbf{r}^N) = \sum_i \phi(\mathbf{r}_i)$. For simplicity, we assume that the intermolecular potential be pairwise additive[1], as will be the case for the particular system that we are interested in. Whether an external potential is present has important implications for the system, as a non-zero potential leads to inhomogeneous equilibrium density profiles. Further, some modifications to the formalism are needed. We start the exposition by assuming no external potential at first and later generalize to the inhomogeneous case with $\Phi \neq 0$.

We are mainly interested in two ensembles, the canonical and the grand canonical ensemble. In the first, the total particle number is fixed, while in the latter, it is variable, controlled by the chemical potential. In both ensembles it is possible to exchange energy with the external reservoir. The temperature of the reservoir will be denoted by T, the inverse temperature by $\beta = 1/k_B T$ and the Boltzmann constant by k_B.

[1] That means there are no explicit three-body interactions.

Canonical ensemble As we are confronted with systems that consist of large numbers of particles and thermal fluctuations, in order to make predictions we need an expression for the probability of certain configurations of the system. It is given in the canonical ensemble by the equilibrium probability distributions $f_0^{[N]}(\mathbf{r}^N, \mathbf{p}^N)$, with the particle number N as index, because it is fixed. It turns out that the probability of a configuration is proportional to the Boltzmann factor $e^{-\beta\mathcal{H}}$. We therefore define a normalization constant as

$$Q_N = \frac{1}{N!h^{3N}} \int e^{-\beta\mathcal{H}} \, \mathrm{d}\mathbf{r}_1 \ldots \mathrm{d}\mathbf{r}_N \, \mathrm{d}\mathbf{p}_1 \ldots \mathrm{d}\mathbf{p}_N \tag{2.2}$$

$$= \frac{1}{\Lambda^{3N} N!} Z_N \tag{2.3}$$

the so-called canonical partition function[2]. Knowing the kinetic energy from Equation (2.1), we can perform the integration over the momenta and factorize out the configurational integral

$$Z_N = \int \mathrm{d}\mathbf{r}_1 \ldots \mathrm{d}\mathbf{r}_N e^{-\beta U_N} \tag{2.4}$$

The momentum integral together with the factor h^{-3N} is absorbed into the thermal wavelength $\Lambda = \left(2\pi m k_B T/h^2\right)^{-1/2}$. We can now write down the definition of the equilibrium probability density as

$$f_0^{[N]}(\mathbf{r}^N, \mathbf{p}^N) = \frac{1}{h^{3N} N!} \frac{\exp(-\beta\mathcal{H})}{Q_N} \tag{2.5}$$

Expectation values of observables are calculated according to

$$\langle B \rangle = \int B(\mathbf{r}^N, \mathbf{p}^N) f_0^{[N]}(\mathbf{r}^N, \mathbf{p}^N) \, \mathrm{d}\mathbf{r}^N \, \mathrm{d}\mathbf{p}^N \tag{2.6}$$

In order to save space we introduce the abbreviation $\mathrm{d}\mathbf{r}_1 \ldots \mathrm{d}\mathbf{r}_N \, \mathrm{d}\mathbf{p}_1 \ldots \mathrm{d}\mathbf{p}_N \equiv \mathrm{d}\mathbf{r}^N \, \mathrm{d}\mathbf{p}^N$ and the (canonical) trace

$$\mathrm{Tr}_\mathrm{c} = \int \cdots \mathrm{d}\mathbf{r}^N \, \mathrm{d}\mathbf{p}^N \tag{2.7}$$

Expectation values can then be written as

$$\langle B \rangle = \mathrm{Tr}_\mathrm{c} \, f_0 B \tag{2.8}$$

[2]The prefactor $N!$ accounts for the indistinguishability of the particles, whereas the factor h^{3N} can be viewed as a necessary consequence of the uncertainty relation. The dimension of h^{3N} is inverse to that of the phase space differentials, the partition function hence dimensionless.

This was the microscopic picture. The connection with the thermodynamic (macroscopic) picture is made by the Helmholtz free energy, defined by

$$F = U - TS \tag{2.9}$$
$$= \mathrm{Tr_c}\, f_0 \left[\mathcal{H} + k_B T \ln f_0\right] \tag{2.10}$$

where we made use of the definition of the entropy

$$S = k_B \,\mathrm{Tr_c}\, f_0 \ln f_0 \tag{2.11}$$

Knowledge of the free energy F allows us to compute equilibrium quantities of the system like the pressure P and the chemical potential μ

$$P = -\left(\frac{\partial F}{\partial V}\right)_{T,N}, \qquad \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \tag{2.12}$$

Grand canonical ensemble Connecting the system to an external particle reservoir allows us to exchange particles with it, making the particle number a variable quantity controlled by the chemical potential μ. It specifies how much energy needs to be brought up in order to add one more particle into the system. The Boltzmann factor changes accordingly and reads now $e^{-\beta(\mathcal{H}-N\mu)}$. The grand canonical partition function is in turn defined by

$$\Xi = \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{N! h^{3N}} \int e^{-\beta\mathcal{H}}\, \mathrm{d}\mathbf{r}_1 \ldots \mathrm{d}\mathbf{r}_N\, \mathrm{d}\mathbf{p}_1 \ldots \mathrm{d}\mathbf{p}_N \tag{2.13}$$
$$= \sum_{N=0}^{\infty} e^{\beta\mu N} Q_N = \sum_{N=0}^{\infty} \frac{z^N}{N!} Z_N \tag{2.14}$$

where now we need to sum over all possible numbers of particles that could be present in the system and their corresponding canonical Boltzmann factor. It is common to define the so-called activity

$$z = \frac{\exp(\beta\mu)}{\Lambda^3} \tag{2.15}$$

The equilibrium probability density reads in the grand canonical case[3]

$$f_0(\mathbf{r}^N, \mathbf{p}^N, N) = \frac{\exp\left[-\beta(\mathcal{H} - N\mu)\right]}{\Xi} \tag{2.16}$$

[3] It might seem that there is an inconsistency between the definition of Equation (2.5) and Equation (2.16), regarding the prefactors. It is however correct if we take into account the changed definition of the grand canonical trace which now includes the factor h^{-3N}.

the corresponding expectation values are given by

$$\begin{aligned}\langle B\rangle &= \sum_{N=0}^{\infty} \frac{1}{h^{3N}N!} \int B(\mathbf{r}^N, \mathbf{p}^N) f_0(\mathbf{r}^N, \mathbf{p}^N, N)\, \mathrm{d}\mathbf{r}^N\, \mathrm{d}\mathbf{p}^N \\ &= \mathrm{Tr}_{\mathrm{gc}}\, f_0 B \end{aligned} \tag{2.17}$$

with the grand canonical trace containing a sum over all possible particle numbers

$$\mathrm{Tr}_{\mathrm{gc}} = \sum_{N=0}^{\infty} \frac{1}{h^{3N}N!} \int \cdots \mathrm{d}\mathbf{r}^N\, \mathrm{d}\mathbf{p}^N \tag{2.18}$$

The corresponding thermodynamic potential is the grand potential Ω given by

$$\Omega = F - N\mu \tag{2.19}$$
$$= k_B T \ln \Xi \tag{2.20}$$
$$= -PV \tag{2.21}$$

where P stands for the pressure. Note that both probability distributions are connected by the relation

$$\frac{1}{h^{3N}N!} f_0(\mathbf{r}^N, \mathbf{p}^N, N) = p(N) f_0^{[N]}(\mathbf{r}^N, \mathbf{p}^N) \tag{2.22}$$

where

$$p(N) = \frac{1}{\Xi} \frac{z^N}{N!} Z^N \tag{2.23}$$

The probability distribution in the GC ensemble can therefore be thought of as a weighted sum of probabilities with fixed particle number N in the canonical ensemble.

Density distribution functions One of the central observable is the n-particle distribution function, which in the canonical ensemble, is given by

$$\rho_N^{(n)}(\mathbf{r}^n) = \frac{N!}{(N-n)!} \frac{1}{Z_N} \int e^{-\beta U_N(\mathbf{r}^N)}\, \mathrm{d}\mathbf{r}_{n+1} \ldots \mathrm{d}\mathbf{r}_N \tag{2.24}$$

where the factorial factors in front are needed because there are N possible ways (finding n out of possible N particles) to choose the particle at position $\mathbf{r}_1$, $N-1$ for the particle at $\mathbf{r}_2$, and so on. The corresponding quantity in the GC ensemble is

$$\rho^{(n)}(\mathbf{r}^n) = \sum_{N=n}^{\infty} p(N) \rho_N^{(n)}(\mathbf{r}^n) \tag{2.25}$$

with $p(N)$ again given by Equation (2.23). Most often however, we are concerned with the one- and two-body density distribution. They can be written as expectation value in the grand canonical ensemble as

$$\rho^{(1)}(\mathbf{r}) = \langle \hat{\rho}(\mathbf{r}) \rangle = \left\langle \sum_{i=0}^{N} \delta(\mathbf{r} - \mathbf{r}_i) \right\rangle \tag{2.26}$$

$$\rho^{(2)}(\mathbf{r}, \mathbf{r}') = \left\langle \sum_{i \neq j}^{N} \delta(\mathbf{r} - \mathbf{r}_i)\delta(\mathbf{r}' - \mathbf{r}_j) \right\rangle \tag{2.27}$$

Since the number of particles is not fixed, the integral over the one-particle density yields the expectation value of number of particles:

$$\int \rho^{(1)}(\mathbf{r}_1)\, \mathrm{d}\mathbf{r}_1 = \langle N \rangle \tag{2.28}$$

with the second-order particle density

$$\int \rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)\, \mathrm{d}\mathbf{r}_1\, \mathrm{d}\mathbf{r}_2 = \left\langle N^2 \right\rangle - \langle N \rangle \tag{2.29}$$

and so on. The equivalent result in the canonical ensemble is given by

$$\int \rho_N^{(1)}(\mathbf{r}_1)\, \mathrm{d}\mathbf{r}_1 = N \tag{2.30}$$

Comparing both ensembles in the limit of a non-interacting, i.e. ideal gas, we observe the following difference

$$\rho^{(2)} = \rho^2 \leftrightarrow \rho_N^{(2)} = \rho^2 \left(1 - \frac{1}{N}\right) \tag{2.31}$$

This fact has implications on some of the results determined from Monte Carlo simulations. The radial distribution function that will be introduced in the next section has a slightly different value for the long-range asymptote ($r \to \infty$), depending on the ensemble that is used in the simulation.

2.1.1 Modifications due to angular degrees of freedom

The above exposition only used spatial degrees of freedom in the phase space integrations. Since we are treating an anisotropic system, the phase space differential should also include the angular variables $\Omega = (\phi, \theta, \chi)$ and the corresponding

canonical momentum p_Ω. Here we make use of the Euler angle parametrization for the orientational configuration of the particle. More details on this parametrization will be given later. Since in the context of the formalism, these coordinates are just generalized coordinates the differential for Ω reads $\mathrm{d}\Omega' = \mathrm{d}\phi\,\mathrm{d}\theta\,\mathrm{d}\chi$, which differs by a factor $\sin\theta$ from the differential in the context of integration over SO(3), hence the use of the prime. The Hamiltonian now includes the anisotropic interparticle potential $U(\mathbf{r}^N, \Omega^N)$ and the rotational kinetic energy, given by [GG84] (Chapter 3 and corresponding Appendix)

$$K_r = \sum_{i=1}^{N} \sum_{\alpha=x,y,z} J_{i\alpha}^2 / 2I_\alpha \tag{2.32}$$

with the angular momentum J_i of particle i, the inertial tensor I, and its diagonal elements I_α. The patchy particle being a spherical top molecule has $I_x = I_y = I_z \equiv I$. In order to perform the phase space integral we need to express the angular momentum appearing in the kinetic energy with the generalized angular momenta p_Ω. This transformation can be found, however it mixes the coordinates Ω and p_Ω as follows

$$J_x = -p_\phi \csc\theta \cos\chi + p_\theta \sin\chi + p_\chi \cot\theta \cos\chi \tag{2.33}$$

$$J_y = p_\phi \csc\theta \sin\chi + p_\theta \cos\chi - p_\chi \cot\theta \sin\chi \tag{2.34}$$

$$J_z = p_\chi \tag{2.35}$$

As a consequence we cannot *directly* factorize the probabilities

$$P(\mathbf{r}^N, \mathbf{p}^N, \Omega^N, p_\Omega^N) \neq P(\mathbf{p}^N) P(p_\Omega^N) P(\mathbf{r}^N, \Omega^N) \tag{2.36}$$

as the terms depending on the coordinates p_Ω^N and Ω^N cannot be separated. In order to make progress, one has two choices, either keep p_Ω as integration variable and rewrite K_r in terms of p_Ω, or change the integration variable from p_Ω to $\mathbf{J}$ and get a new probability distribution $P(\mathbf{r}^N, \mathbf{p}^N, \Omega^N, \mathbf{J}^N)$. Both routes are possible, though the second is more straightforward mathematically. The result is that the probability distribution factorizes into configurational, translational and rotational parts

$$P(\mathbf{r}^N, \mathbf{p}^N, \Omega^N, \mathbf{J}^N) = P(\mathbf{r}^N, \mathbf{p}^N) P(\Omega^N) P(\mathbf{J}^N) \tag{2.37}$$

The latter two can be absorbed into the thermal wavelength, leaving us only with the configurational description of the system.

2.2 Integral equation theories

The n-particle correlation function is defined in either ensemble (canonical or grand canonical) by the n-particle density distribution function

$$g(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \frac{\rho^{(n)}(\mathbf{r}_1, \ldots, \mathbf{r}_n)}{\rho^{(1)}(\mathbf{r}_1) \cdots \rho^{(1)}(\mathbf{r}_n)} \tag{2.38}$$

In the case of $n = 2$ it becomes the pair distribution function

$$g(\mathbf{r}_1, \mathbf{r}_2) = \frac{\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)}{\rho^{(1)}(\mathbf{r}_1)\rho^{(1)}(\mathbf{r}_2)} \tag{2.39}$$

When the system is homogeneous and the interaction between the particles is isotropic, then g will only depend on the scalar radial variable r, usually called the radial distribution function (RDF)

$$g(r) = \frac{\rho^{(2)}(r)}{\rho^2} \tag{2.40}$$

with the bulk density ρ, which gives us information about the local density distribution around a test particle. Apart from this, it can also be used to infer properties like the pressure (over the so-called virial route)

$$\frac{\beta P}{\rho} = 1 - \frac{\rho\beta}{6} \int \frac{\mathrm{d}u(r)}{\mathrm{d}r} g(\mathbf{r}) r \, \mathrm{d}\mathbf{r} \tag{2.41}$$

The total correlation function is defined by

$$h(\mathbf{r}_1, \mathbf{r}_2) = g(\mathbf{r}_1, \mathbf{r}_2) - 1 \tag{2.42}$$

and results from subtracting the "constant (ideal) background"[4] from g, such that for large separations h vanishes. One can now, given the physical nature of the correlation function, propose a direct correlation function $c(\mathbf{r})$ as follows

$$c(\mathbf{r}) = h(\mathbf{r}) - \rho \int c(\mathbf{r}')h(|\mathbf{r} - \mathbf{r}'|) \, \mathrm{d}\mathbf{r}' \tag{2.43}$$

the so-called *Ornstein–Zernike* (OZ) equation. Its physical interpretation is that the total correlation between two points is the sum of the direct correlation between them plus a term that repeats this decomposition for a third point, which is integrated over. In Section 2.4 we will present an alternative way to define the direct

[4] For the ideal gas $g(r) = 1$.

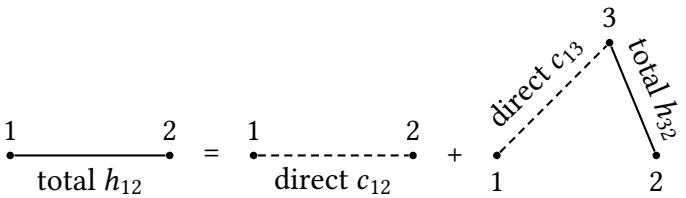


Figure 2.1: Graphical representation of Equation (2.43). The vector $\mathbf{r}$ corresponds to the distance between points (1) and (2), while the integration variable $\mathbf{r}'$ corresponds to the point (3).

correlation function based on density functional theory. The equivalence of both definitions is shown in [HM13]. To complete the taxonomy of correlation functions we can introduce the indirect correlation function $\gamma(\mathbf{r})$ as

$$\gamma(\mathbf{r}) = h(\mathbf{r}) - c(\mathbf{r}) \tag{2.44}$$

$$= \rho \int c(\mathbf{r}')h(|\mathbf{r} - \mathbf{r}'|)\,\mathrm{d}\mathbf{r}' \tag{2.45}$$

2.2.1 Closure relations

With only the Ornstein–Zernike equation, the system $\{c(\mathbf{r}), g(\mathbf{r})\}$ is underdetermined. In order to solve it, is necessary to specify a further constraint, the so-called "closure relation". The general form of the closure relation is given by

$$c(\mathbf{r}) = h(\mathbf{r}) - \ln y(\mathbf{r}) + B(\mathbf{r}) \tag{2.46}$$

$$= f(\mathbf{r})y(\mathbf{r}) + t(r) \tag{2.47}$$

or alternatively

$$h(\mathbf{r}) + 1 = \exp\left(-\beta u(\mathbf{r}) + \gamma(\mathbf{r}) + B(\mathbf{r})\right) \tag{2.48}$$

The quantities that appear are

- the bridge function $B(\mathbf{r})$
- the cavity function
$$y(\mathbf{r}) = g(\mathbf{r})e^{\beta u(\mathbf{r})} \tag{2.49}$$
- the tail function $t(\mathbf{r})$
- the Mayer f-function
$$f(\mathbf{r}) = e^{-\beta u(\mathbf{r})} - 1 \tag{2.50}$$

Two of the most popular closure are now discussed.

Percus–Yevick closure It is given by

$$B(\mathbf{r}) = \ln y(\mathbf{r}) - y(\mathbf{r}) + 1 \tag{2.51}$$
$$c(\mathbf{r}) = f(\mathbf{r})y(\mathbf{r}) \tag{2.52}$$

which means that it assumes $t(\mathbf{r}) = 0$. Usually the resulting system of equations is solved numerically but in the case of hard spheres and sticky hard spheres[5] analytical solutions do exist.

Hypernetted-chain closure This closure results from setting the bridge function to zero

$$B(\mathbf{r}) = 0 \tag{2.53}$$
$$c(\mathbf{r}) = h(\mathbf{r}) - \ln y(\mathbf{r}) \tag{2.54}$$

In general one cannot say whether one closure is more accurate than the other, as this depends on the interparticle potential. Judging by diagrammatic methods it may seem that the HNC approximation includes more terms, however due to cancellation of certain terms it turns out that in the case of hard-core interaction the PY closure gives more accurate results, see [San16].

2.3 Inhomogeneous systems

In the previous expositions we only considered homogeneous systems, as there was no external potential or phase boundaries. By introducing a position-dependent potential, the Hamiltonian gains an additional term, modifying some of the previous equations. Using the δ-function description of the particle density we can write

$$\Phi_N(\mathbf{r}^N) = \sum_{i=1}^{N} \phi(\mathbf{r}_i) = \int \mathrm{d}\mathbf{r} \rho(\mathbf{r}) \phi(\mathbf{r}) \tag{2.55}$$

Since the external potential only couples to the positional degrees of freedom (and not the generalized momenta) only the configurational integral is modified

$$Z_N = \int \mathrm{d}\mathbf{r}^N e^{-\beta U_N} \to \int \mathrm{d}\mathbf{r}^N e^{-\beta U_N} \left(\prod_{i=1}^{N} \exp\left[-\beta \phi(\mathbf{r}_i)\right] \right) \tag{2.56}$$

[5] Sticky hard spheres are hard spheres with an additional attractive interaction when two such spheres touch. The sharp positional dependence of the interaction (δ-function) makes analytical computations easier, however solving the PY closure remains a challenging task.

The grand partition function becomes

$$\Xi = \sum_{N=0}^{\infty} \frac{1}{N!} \int e^{-\beta U_N} \left(\prod_{i=1}^{N} z \exp\left[-\beta\phi(\mathbf{r}_i)\right] \right) \mathrm{d}\mathbf{r}^N \tag{2.57}$$

Separating the activity into the thermal wave length and chemical potential, one can rewrite the previous equation as

$$\Xi = \sum_{N=0}^{\infty} \frac{1}{N!} \int e^{-\beta U_N} \left(\prod_{i=1}^{N} \Lambda^{-3} \exp\left[\beta\psi(\mathbf{r}_i)\right] \right) \mathrm{d}\mathbf{r}^N \tag{2.58}$$

with the intrinsic chemical potential

$$\psi(\mathbf{r}) = \mu - \phi(\mathbf{r}) \tag{2.59}$$

The partition function is then formally equal to the homogeneous one, except for the replacement of $\mu \rightarrow \psi(\mathbf{r})$. One can also rewrite the thermodynamic potentials for the inhomogeneous case. We can think of the external potential ϕ replacing the thermodynamic variable V by shifting the available volume into the function ϕ. Then, the change in internal energy $U(S, V, N)$ becomes

$$\delta U = T\,\delta S + \int \rho^{(1)}(\mathbf{r})\,\delta\phi(\mathbf{r})\,\mathrm{d}\mathbf{r} + \mu\,\delta N \tag{2.60}$$

The change in Helmholtz free energy $F = U - TS$ is consequently

$$\delta F = -S\,\delta T + \int \rho^{(1)}(\mathbf{r})\,\delta\phi(\mathbf{r})\,\mathrm{d}\mathbf{r} + \mu\,\delta N \tag{2.61}$$

In analogy to the intrinsic chemical potential[6] in Equation (2.59) we define the intrinsic free energy

$$\mathcal{F} = F - \int \rho^{(1)}(\mathbf{r})\phi(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{2.62}$$

with the differential being

$$\delta\mathcal{F} = -S\,\delta T - \int \delta\rho^{(1)}(\mathbf{r})\phi(\mathbf{r})\,\mathrm{d}\mathbf{r} + \mu\,\delta N \tag{2.63}$$

$$= -S\,\delta T + \int \delta\rho^{(1)}(\mathbf{r})\psi(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{2.64}$$

[6] For the intrinsic chemical potential we combined the chemical and the external potential.

where $\psi(\mathbf{r})$ appears as conjugate variable to $\delta\rho^{(1)}(\mathbf{r})$. For the grand potential $\Omega = F - N\mu$, it follows that

$$\delta\Omega = -S\,\delta T - \int \rho^{(1)}(\mathbf{r})\,\delta\psi(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{2.65}$$

By inspection of Equations (2.61) and (2.65) and the fact that we replaced the variable V by the function $\phi(\mathbf{r})$ it is clear that now we need to treat the grand potential and the intrinsic free energy as functionals $\Omega[\psi(\mathbf{r})]$ and $\mathcal{F}[\rho(\mathbf{r})]$ respectively. By computing

$$\frac{\delta\mathcal{F}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} = \psi(\mathbf{r}) \tag{2.66}$$

and

$$\frac{\delta\Omega[\psi(\mathbf{r})]}{\delta\psi(\mathbf{r})} = -\rho(\mathbf{r}) \tag{2.67}$$

we see that they are related by a generalized Legendre transformation. From the definition $\Omega = F - N\mu$ it follows that

$$\Omega[\psi] = \mathcal{F}[\rho] - \int \rho^{(1)}\psi(\mathbf{r}) \tag{2.68}$$

Finally, we show that $\mathcal{F}$ can be written also as an ensemble average. For this, we use the definition of $f_0(\mathbf{r}^N, \mathbf{p}^n, N)$ and that of the Hamiltonian in Equation (2.1) to get

$$\ln f_0 = -\beta\mathcal{H} + \beta\mu N - \ln\Xi \tag{2.69}$$

$$= -\beta K_N - \beta U_N - \beta\Phi_N + N\beta\mu \tag{2.70}$$

Taking the average and using $\beta\Omega = \ln\Xi$

$$\langle K_N + U_N + k_B T \ln f_0\rangle = \Omega + \int \mathrm{d}\mathbf{r}\rho^{(1)}(\mathbf{r})\,(\mu - \phi) \tag{2.71}$$

$$= \Omega + \int \mathrm{d}\mathbf{r}\rho^{(1)}(\mathbf{r})\psi(\mathbf{r}) \tag{2.72}$$

$$= \mathcal{F} \tag{2.73}$$

2.4 Density functional theory (DFT)

The crucial ingredient for arriving at a useful method is parametrizing the free energy functional with the particle density, the main point of density functional

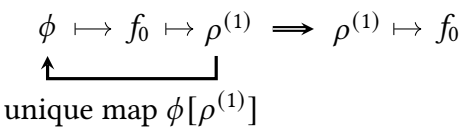


Figure 2.2: Diagram showing the order of mappings between the external potential, the equilibrium phase space probability and the particle distribution function. The arrows stand for the existence of a functional mapping.

theory. While we know from Section 2.3, that the one-particle distribution function $\rho^{(1)}(\mathbf{r})$ is in general a functional of the external potential, it can indeed be proved that there exist a unique mapping between density profile and external potential [Eva79]. In the following we give a sketch of the underlying principle. From the definition of f_0 in Equation (2.16), together with Equation (2.1), we know that it is a functional of the external potential $\phi(\mathbf{r})$. We can write the density distribution as

$$\rho^{(1)} = \mathrm{Tr}_{\mathrm{gc}}\left[f_0 \hat{\rho}(\mathbf{r})\right] \tag{2.74}$$

hence $\rho^{(1)}$ is a functional of f_0, and therefore also of $\phi(\mathbf{r})$. The situation is visualized in Figure 2.2. One can further show that the mapping $\phi(\mathbf{r}) \to \rho^{(1)}$ is unique (see Appendix B of [HM13]), hence we can replace the functional dependence of f_0 as

$$f_0[\phi(\mathbf{r})] \to f_0[\rho^{(1)}(\mathbf{r})] \tag{2.75}$$

The argument is completed by the fact that we can write the intrinsic free energy as trace over f_0, i.e.

$$\begin{aligned}
\mathcal{F}[f_0] &= \mathrm{Tr}\left[f_0\left(K_N + U_N + k_B T \ln f_0\right)\right] && (2.76)\\
&= \mathrm{Tr}\left[f_0\left(K_N + k_B T \ln f_0\right)\right] + \mathrm{Tr}\left[f_0 U_N\right] && (2.77)\\
&= \mathcal{F}_{\mathrm{id}}[\rho^{(1)}] + \mathcal{F}_{\mathrm{ex}}[\rho^{(1)}] && (2.78)\\
&= \mathcal{F}[\rho^{(1)}] && (2.79)
\end{aligned}$$

Similarly one can also write the grand potential as a functional of the one-body density

$$\begin{aligned}
\Xi[f_0] &= \mathrm{Tr}\left[f_0\left(\mathcal{H} - \mu N + k_B T \ln f_0\right)\right] && (2.80)\\
&= \mathcal{F}[\rho^{(1)}] + \mathrm{Tr}\left[f_0\left(\Phi_N - \mu N\right)\right] && (2.81)\\
&= \mathcal{F}[\rho^{(1)}] + \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\left(\phi(\mathbf{r}) - \mu\right) && (2.82)\\
&= \Xi[\rho^{(1)}] && (2.83)
\end{aligned}$$

In the literature, $\phi(\mathbf{r})$ is frequently denoted by $V^{\text{ext}}(\mathbf{r})$, which we will adopt in the following. It can further be shown, that the equilibrium density distribution $\rho(\mathbf{r})$ minimizes the grand potential, meaning that for every other density distribution $\rho'(\mathbf{r}) \neq \rho(\mathbf{r})$: $\Xi[\rho'] > \Xi[\rho] = \Xi$. Written differently, this leads to the main equation of classical density functional theory

$$\left.\frac{\delta\Xi[\rho]}{\delta\rho(\mathbf{r})}\right|_{\rho=\rho_{\text{eq}}} = \left.\frac{\delta\mathcal{F}[\rho]}{\delta\rho(\mathbf{r})}\right|_{\rho=\rho_{\text{eq}}} + V^{\text{ext}}(\mathbf{r}) - \mu = 0 \tag{2.84}$$

It is common to split the free energy into ideal-gas term and the "excess"[7] term that it is due to interactions: $\mathcal{F}[\rho] = \mathcal{F}_{\text{id}}[\rho] + \mathcal{F}_{\text{ex}}[\rho]$, where

$$\beta\mathcal{F}_{\text{id}}[\rho] = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r}) \left[\ln(\rho(\mathbf{r})\Lambda^3) - 1\right] \tag{2.85}$$

We can now solve Equation (2.84) for the density profile $\rho(\mathbf{r})$, giving us the *implicit*[8] equation

$$\rho(\mathbf{r}) = \exp\left(-\beta V^{\text{ext}}(\mathbf{r}) + c^{(1)}(\mathbf{r}) + \beta\mu\right) \tag{2.86}$$

$$= \rho_b \exp\left(-\beta V^{\text{ext}}(\mathbf{r}) + c^{(1)}(\mathbf{r}) + \beta\mu_{\text{ex}}\right) \tag{2.87}$$

where $c^{(1)}$ is the one-body direct correlation function (DCF), defined by

$$c^{(1)}(\mathbf{r}) = -\frac{\delta\beta\mathcal{F}_{\text{ex}}[\rho]}{\delta\rho(\mathbf{r})} \tag{2.88}$$

We can also define the general n-body direct correlation function

$$c^{(n)}(\mathbf{r}_1, \ldots, \mathbf{r}_n) = -\frac{\delta^n\beta\mathcal{F}_{\text{ex}}[\rho]}{\delta\rho(\mathbf{r}_1)\cdots\delta\rho(\mathbf{r}_n)} \tag{2.89}$$

For $n = 2$ and evaluated at the bulk density ρ_b we get the two-body DCF defined in Equation (2.43).

2.4.1 Density functional approximations

Since the (excess) free energy functional is in general unknown and a direct computation is not possible, it is necessary to employ approximations.

[7] Meaning in excess of the ideal gas term.

[8] The density profile $\rho(\mathbf{r})$ enters into $c^{(1)}$, therefore appearing on both sides of the equation, making it implicit.

Local density approximation (LDA) For this approximation one starts with a known expression for the free energy density per particle $f(\rho)$ and assumes the following form

$$\beta\mathcal{F}_{\text{ex}}[\rho] = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\beta f(\rho(\mathbf{r})) \tag{2.90}$$

i.e. the functional is an integral over the bulk free energy density evaluated at the corresponding (local) density. Though simple, the approximation is not very accurate and does not find much application, except for calculations concerning liquid–vapor interfaces, where there is only weak short-range structure. It may however be seen as the starting point for the next family of approximations which is indeed much more accurate.

Weighted density approximation (WDA) As we have noted, except for simple cases, the LDA cannot be used to accurately model liquids, due to the assumption of locality. In order to incorporate non-locality, one can "coarse grain" the density profile over a limited region using a convolution with a kernel or *weight function* $w(r)$. The resulting weighted density

$$\overline{\rho}(\mathbf{r}) = \int \mathrm{d}\mathbf{r}'\rho(\mathbf{r}')w(\mathbf{r}-\mathbf{r}') \tag{2.91}$$

replaces the local density from before and the free energy per particle is replaced by the function[9] $\Phi(\overline{\rho})$

$$\beta\mathcal{F}_{\text{ex}}[\rho] = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\beta\Phi(\overline{\rho}(\mathbf{r})) \tag{2.92}$$

There are different strategies to determine Φ, like that of Tarazona [Tar85].

Mean-field approximation A further possibility is to expand the exact free energy up to second order in a Taylor series. One starts with separating the full interparticle potential into a reference part u_0 and a perturbation w, which is controlled by a parameters λ, as in

$$u_\lambda(r) = u_0(r) + \lambda w(r) \tag{2.93}$$

[9] Note that this quantity is not related to the Φ_N that appeared in the many-body Hamiltonian.

The excess free energy corresponding to u_0 is given by $\mathcal{F}_{\text{ex}}^0$, while the full excess free energy reads

$$\beta\mathcal{F}_{\text{ex}}[\rho] = \mathcal{F}_{\text{ex}}^0[\rho] + \frac{1}{2}\int_0^1 \mathrm{d}\lambda \int \rho^{(2)}(\mathbf{r},\mathbf{r}';\lambda) w(\mathbf{r},\mathbf{r}')\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \tag{2.94}$$

$$= \mathcal{F}_{\text{ex}}^0[\rho] + \frac{1}{2}\int \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\rho(\mathbf{r})\rho(\mathbf{r}')w(\mathbf{r}-\mathbf{r}') + \mathcal{F}_{\text{corr}}[\rho] \tag{2.95}$$

where the correlation free energy is given by

$$\mathcal{F}_{\text{corr}}[\rho] = \int_0^1 \mathrm{d}\lambda \int \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\rho(\mathbf{r})\rho(\mathbf{r}')h(\mathbf{r},\mathbf{r}';\lambda)w(\mathbf{r}-\mathbf{r}') \tag{2.96}$$

By neglecting this correlation term we arrive at the mean-field ansatz. It is easy to see that its validity depends on the strength of the correlation h, which is weakest for low densities and high temperatures.

2.4.2 Fundamental measure theory

The approximations presented above represent perturbative expansions and as such have only a limited range of validity (i.e. slowly varying profiles or low density systems). An exception to this, is a family of functionals for hard sphere systems, under the name of fundamental measure theory. Its origin goes back to Y. Rosenfeld, who derived a functional by making a general ansatz using scaled particle theory and determining unknowns by dimensional analysis [Ros89]. The name fundamental measure refers to the weight functions (analogous to the weights used in the WDA) that represent geometrical properties of spheres in different dimensions (diameter, surface, volume). Over the years more advanced functionals have been developed using the same or similar weight functions, achieving very high prediction accuracy when compared to simulations.

In our study we use the Rosenfeld version of FMT, which is numerically easy to implement and sufficiently accurate for the densities we are interested in. The DCF calculated according to Equation (2.89) is equal to the one resulting from the PY closure and the integral equation route. Note that the Rosenfeld functional is not able to predict crystallization, a shortcoming that is remedied by tensor functionals starting with [Tar00] and later used in the family of White Bear FMT functionals [Rot+02; HR06]. The principal weight functions are $w_2(\mathbf{r})$, $w_3(\mathbf{r})$ and $\mathbf{w}_2(\mathbf{r})$, while

the others are only scaled variants of these. They are given by

$$w_2(\mathbf{r}) = \delta(R - |\mathbf{r}|) \tag{2.97}$$
$$w_3(\mathbf{r}) = \Theta(R - |\mathbf{r}|) \tag{2.98}$$
$$\mathbf{w}_2(\mathbf{r}) = \frac{\mathbf{r}}{r} w_2(\mathbf{r}) \tag{2.99}$$

The scaled variants are

$$w_0(\mathbf{r}) = \frac{w_2(\mathbf{r})}{4\pi R^2} \tag{2.100}$$
$$w_1(\mathbf{r}) = \frac{w_2(\mathbf{r})}{4\pi R} \tag{2.101}$$
$$\mathbf{w}_1(\mathbf{r}) = \frac{\mathbf{w}_2(\mathbf{r})}{4\pi R} \tag{2.102}$$

The weighted densities result from convolutions of the particle density with these weight functions

$$n_\nu(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \rho(\mathbf{r}') w_\nu(\mathbf{r} - \mathbf{r}') = (\rho * w_\nu)(\mathbf{r}) \tag{2.103}$$

and similarly for the vector weight functions. The functional has the following form

$$\beta\mathcal{F}_{\mathrm{hs}}[\rho] = \int \mathrm{d}\mathbf{r}\, \Phi_{\mathrm{hs}}(\{n_\nu(\mathbf{r})\}) \tag{2.104}$$

$$\Phi_{\mathrm{hs}}(\{n_\nu(\mathbf{r})\}) = -n_0 \ln(1 - n_3) + \frac{n_1 n_2 - \mathbf{n}_1 \cdot \mathbf{n}_2}{1 - n_3} + \frac{n_2^3 - 3 n_2 \mathbf{n}_2 \cdot \mathbf{n}_2}{24\pi(1 - n_3)^2} \tag{2.105}$$

2.4.3 Calculating density profiles

We have seen that the equilibrium density profiles are determined by the implicit equation given in Equation (2.87). Since it is implicit it can't be solved directly. Instead, we use iterative methods to find a solution.

The usual procedure is as follows

- Initialize the quantities that do not change: $V^{\mathrm{ext}}(\mathbf{r})$, μ_{ex}, ρ_b
- Depending on the excess free energy functional that is used, prepare the function $c^{(1)} = -\frac{\delta\beta\mathcal{F}_{\mathrm{ex}}}{\delta\rho(\mathbf{r})}$, which given a density profile $\rho(\mathbf{r})$ gives the corresponding one-body DCF. Note that $c^{(1)}[\rho_b](\mathbf{r}) = -\mu_{\mathrm{ex}} = \text{const.}$, which needs to be the case because the density $\rho(\mathbf{r})$ at a point that is far from the applied external potential and hence "feels" no influence, must be equal to the bulk density

- Initialize the first trial density profile by some value, e.g. $\rho_{i=0}(\mathbf{r}) = \rho_b \forall \mathbf{r}$ and compute the r.h.s. of the implicit equation resulting in the output density $\overline{\rho}_i(\mathbf{r})$
- Compute a new trial solution by mixing the previous input and output densities

$$\rho_{i+1}(\mathbf{r}) = (1-\alpha)\rho_i(\mathbf{r}) + \alpha\overline{\rho_i}(\mathbf{r}) \tag{2.106}$$

  the so-called Picard iteration. The mixing parameter α stabilizes the procedure and helps to prevent density profiles that are unphysical (impossible packing fractions), leading to diverging terms in the $c^{(1)}$ computation. It should be chosen small, but not too small as this hinders fast convergence.
- Repeat until convergence criterion is met, i.e. $||\rho_{i+1}(\mathbf{r}) - \rho_i(\mathbf{r})|| < \delta$ for some small value of δ.

The Picard iteration scheme is simple and robust, however there exist more advanced schemes that are able to reach convergence faster, i.e. with fewer function evaluations. Our code will make use of Anderson mixing (also known as Anderson acceleration, Pulay mixing or DIIS[10]). The two principal changes are that (i) more than the last "output" values are considered when constructing a new trial solution and (ii) the mixing parameter (now with a vector with more than one dimension) is not fixed but optimized such that the matrix of the previous residuals times the mixing parameter α is minimal, with the condition that $\sum_i \alpha_i = 1$ A more in-depth analysis with comparison to other methods can be found in [Chu+21].

The structure of the FMT functionals can be exploited to speed up the calculation of $c^{(1)}$. Convolutions appear twice during the evaluation of the one-body DCF, once when computing the weighted densities, and once due to the chain-rule differentiation of the weighted densities w.r.t. the density profile

$$c^{(1)}(\mathbf{r}) = -\sum_\nu \int \mathrm{d}\mathbf{r}' \frac{\partial \Phi_{\mathrm{hs}}}{\partial n_\nu(\mathbf{r}')} \frac{\delta n_\nu(\mathbf{r}')}{\delta \rho(\mathbf{r})} \tag{2.107}$$

$$= -\sum_\nu \int \mathrm{d}\mathbf{r}' \frac{\partial \Phi_{\mathrm{hs}}}{\partial n_\nu(\mathbf{r}')} w_\nu(\mathbf{r}' - \mathbf{r}) \tag{2.108}$$

the last line being again a convolution. These convolutions can be done using the discrete Fourier transformation, which can be performed very efficiently using the fast Fourier transform (FFT) algorithm.

[10] DIIS, standing for direct inversion in the iterative subspace.

When the external potential has certain symmetries, for example planar symmetry, such that $V^{\text{ext}}(\mathbf{r}) = V^{\text{ext}}(z)$, we can speed up computations by solving the lower-dimensional problem. In that case we can perform the integrals (for the weighted densities and $c^{(1)}$) over x and y analytically, see Appendix A. It turns out that in the case of planar symmetry the resulting integrals keep their convolutional form and can hence be performed rather efficiently, though some care has to be taken with the vectorial weighted densities, see [Rot10].

2.5 Wigner basis for SO(3)

In the previous section we described the configuration of a particle, or the density distribution with just the spatial variable $\mathbf{r}$. If the particles have further degrees of freedom, these need to be included into the description as well. We are specifically interested in particles that have an anisotropic interaction potential, therefore it is necessary to take into account the rotation of each particle. For the parametrization we chose Euler angles, due to them being relatively easy to visualize and implement. Alternatives are rotation matrices or quaternions, the latter of which are especially well suited for numerical applications as they do not suffer from some of the peculiarities that plague Euler angles. Nevertheless, we settle for the simplest choice we did not encounter any numerical problems and most of the literature uses the same parametrization.

In order to fully describe the orientation of the general body (that means no symmetries) one needs three real numbers, in the Euler angle case, the rotation angles around certain axes. Multiple conventions for how these axes are to be chosen, exist. We employ the one where the particle is rotated around the following axes: *zyz*. Further details, like the use of passive or active transformations are given later. From a mathematical point of view, the rotations form a group called SO(3), the special orthogonal group. Orthogonal, because the matrices representing the rotations fulfill $RR^\top = R^\top R = 1$ and special, because it includes only proper rotations with $\det R = 1$.

The principal problem when treating anisotropic particles is that every observable has in addition to its spatial dependence also an angular dependence which can become difficult to handle. Especially when doing simulations and gathering data the number of bins to fill increase dramatically. For the conservative choice of $\Delta\phi = 0.2$ for the angular bin spacing (for all three angles) we get

$$\frac{(2\pi)^2 2}{(\Delta\phi)^3} \approx 10^4 \tag{2.109}$$

bins to fill, which is in addition to every possible bin in the spatial coordinate. Consequently, statistics is rather poor and long runs are necessary to produce usable data.

Similarly, the high dimensionality of physical quantities, for example the density distribution $\rho(\mathbf{r}, \Omega)$, does not lend itself to physical interpretations readily, even a sensible visualization is already not clear as the dataset has more than three dimensions. In order to simplify the treatment and reduce the dimensionality of the relevant quantities, we get rid of the angular dependence by projecting on a suitable set of basis functions. This set is given by the so-called Wigner D-matrices, defined as

$$D^l_{mn}(\phi, \theta, \chi) = e^{-im\phi} d^l_{mn}(\theta) e^{-in\chi} \tag{2.110}$$

where for every $l \in \mathbb{N}$ the matrix D^l is square with dimensions $(2l+1) \times (2l+1)$, i.e. $m, n \in \{-l, \ldots, l\}$. The small Wigner d-matrix can be defined by

$$\begin{aligned} d^j_{m'm}(\beta) = {} & [(j+m')!(j-m')!(j+m)!(j-m)!]^{\frac{1}{2}} \\ & \times \sum_{s=s_{\min}}^{s_{\max}} \left[\frac{(-1)^{m'-m+s} \left(\cos\frac{\beta}{2}\right)^{2j+m-m'-2s} \left(\sin\frac{\beta}{2}\right)^{m'-m+2s}}{(j+m-s)!s!(m'-m+s)!(j-m'-s)!} \right] \end{aligned} \tag{2.111}$$

where, $s_{\min} = \max(0, m-m')$ and $s_{\max} = \min(j+m, j-m')$, such that the factorials in the denominator are well-defined. The Wigner D-matrices have multiple uses and as such allow different interpretations. One can think of them as generalizations of spherical harmonics, which form a set of complete orthonormal functions on the unit sphere S^2, which, parametrized by the spherical coordinates (ϕ, θ), is a subset of the Euler angle triplet (ϕ, θ, χ). Indeed, it follows with the previous convention that

$$\overline{D^l_{m0}}(\phi, \theta, 0) = \sqrt{\frac{4\pi}{2l+1}} Y^l_m(\theta, \phi) \tag{2.112}$$

where the line over the quantity stands for complex conjugation, i.e. for a complex number z conjugation changes the sign of the imaginary part

$$\overline{z} = \overline{x + iy} = x - iy \tag{2.113}$$

The Wigner D-matrices are also used when rotating spherical harmonics. The spherical harmonic evaluated at the rotated position ω' is written

$$Y^l_{m'}(\omega') = \sum_m D^l_{mm'}(\Omega) Y^l_m(\omega) \tag{2.114}$$

where the rotation between the coordinate systems of ω and ω' is given by Ω.

An important property of the D-matrices is that they form an irreducible basis of SO(3) for every l and that they realize rotations of irreducible spherical tensors T_l. Spherical tensors T_l of rank l are exactly those objects, that transform under rotations the same way as the spherical harmonics Y_m^l, i.e.

$$T_{lm'} = \sum_m D^l_{mm'}(\Omega) T_{lm} \tag{2.115}$$

Tensors of rank 0 are scalars and thus invariant under rotations. In this case the Wigner D-matrix is equal to the identity matrix and the transformation law is trivially fulfilled. Rank 1 tensors are vectors and their transformation under rotations is realized by an orthogonal matrix R with $\det R = 1$, a rotation matrix. While the Wigner D-matrix with $l = 1$ is not identical to the rotation matrix R, they are related by a basis change, namely from the Cartesian to the spherical basis, given by

$$\mathbf{e}_+ = -\frac{1}{\sqrt{2}}\mathbf{e}_x - \frac{i}{\sqrt{2}}\mathbf{e}_y \tag{2.116}$$

$$\mathbf{e}_- = +\frac{1}{\sqrt{2}}\mathbf{e}_x - \frac{i}{\sqrt{2}}\mathbf{e}_y \tag{2.117}$$

$$\mathbf{e}_0 = \mathbf{e}_z \tag{2.118}$$

This is familiar from how spherical harmonics with Y_m^1 transform, namely as vectors in a transformed basis. Tensors of higher rank are for example the quadrupole tensor or the Wigner D-matrices themselves.

Using this set of complete basis we can thus expand any function of Euler angles Ω and possibly other (independent) variables x as

$$\begin{aligned} f(x,\Omega) &= \sum_{lmn} f^l_{mn}(x) D^l_{mn}(\Omega) \\ &= f^0_{00} + \begin{pmatrix} f^1_{\underline{1},\underline{1}}(x) & f^1_{\underline{1},0}(x) & f^1_{\underline{1},1}(x) \\ f^1_{0,\underline{1}}(x) & f^1_{0,0}(x) & f^1_{0,1}(x) \\ f^1_{1,\underline{1}}(x) & f^1_{1,0}(x) & f^1_{1,1}(x) \end{pmatrix} \odot \begin{pmatrix} D^1_{\underline{1},\underline{1}}(\Omega) & D^1_{\underline{1},0}(\Omega) & D^1_{\underline{1},1}(\Omega) \\ D^1_{0,\underline{1}}(\Omega) & D^1_{0,0}(\Omega) & D^1_{0,1}(\Omega) \\ D^1_{1,\underline{1}}(\Omega) & D^1_{1,0}(\Omega) & D^1_{1,1}(\Omega) \end{pmatrix} \\ &\quad + \dots \end{aligned} \tag{2.119}$$

where the dot product has to be understood element-wise between matrix elements of the same index. Here and in the following we use the convention that underlined indices represent the negative value of that index, i.e. $\underline{1} = -1$. The coefficients f^l_{mn} are determined by

$$f^l_{mn}(x) = \frac{2l+1}{K} \int \mathrm{d}\Omega f(x,\Omega) \overline{D^l_{mn}}(\Omega) \tag{2.120}$$

where

$$\mathrm{d}\Omega = \mathrm{d}\phi\, \mathrm{d}\cos\theta\, \mathrm{d}\chi = \mathrm{d}\phi\, \mathrm{d}\theta\, \mathrm{d}\chi \sin\theta \tag{2.121}$$

and

$$\int \mathrm{d}\Omega = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{d}\theta \sin\theta \int_0^{2\pi} \mathrm{d}\phi = 8\pi^2 \tag{2.122}$$

The Wigner D-matrices fulfill the following properties

(a) Orthogonality

$$\int \mathrm{d}\Omega D^{l_1}_{m_1 n_1}(\Omega) \overline{D^{l_2}_{m_2 n_2}}(\Omega) = \frac{8\pi^2}{2l_1 + 1} \delta_{l_1 l_2} \delta_{m_1 m_2} \delta_{n_1 n_2} \tag{2.123}$$

(b) Unitarity

$$\sum_m \overline{D^l_{mn}}(\Omega) D^l_{mn'}(\Omega) = \delta_{nn'} \tag{2.124}$$

$$\sum_n \overline{D^l_{mn}}(\Omega) D^l_{m'n}(\Omega) = \delta_{mm'} \tag{2.125}$$

More properties and useful identifies can be found in [GG84] and in [VMK88].

2.5.1 Numerical evaluation

Numerically evaluating the Wigner D-matrices for arbitrary arguments is not a trivial task. The formula given in Equation (2.111) can in theory be used, but due to the many large factors resulting from the factorial expressions, loss of precision start to sets in already for low values of l. Fortunately there exist alternative methods that allow for a fast and accurate evaluation of the matrices.

Recursive method according to Choi et al.

Here we describe the method proposed in [Cho+99]. It is based on the fact that Wigner D-matrices with upper index l can be written as linear combinations of those with an index $l - 1$. The recursion stops with $l = 1$, as these D-matrices can easily be evaluated by linear transformation of rotation matrices.

The algorithm assumes that the rotation matrix R corresponding to the three Euler angles for which one wishes to compute $D(\phi, \theta, \chi)$ is known. If this is not the

case, and one knows only the angles ϕ, θ, χ, it is easy to compute it by a matrix multiplication of the three individual rotation matrices

$$R(\theta, \phi, \chi) = R_z(\phi) R_y(\theta) R_z(\chi) = \begin{pmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{pmatrix} \tag{2.126}$$

As mentioned earlier, D^1 is related to the rotation matrix R by a change of basis. The real and imaginary part of the Wigner matrix is given by $D^1 = F + iG$ with

$$F = \begin{pmatrix} (R_{yy} + R_{xx})/2 & R_{xz}/\sqrt{2} & (R_{yy} - R_{xx}/2) \\ R_{zx}/\sqrt{2} & R_{zz} & -R_{zx}/\sqrt{2} \\ (R_{yy} - R_{xx})/2 & -R_{xz}/\sqrt{2} & (R_{yy} + R_{xx}/2) \end{pmatrix} \tag{2.127}$$

$$G = \begin{pmatrix} (R_{yx} - R_{xy})/2 & R_{yz}/\sqrt{2} & -(R_{yx} + R_{xy}/2) \\ -R_{zy}/\sqrt{2} & 0 & -R_{zy}/\sqrt{2} \\ (R_{yx} + R_{xy})/2 & -R_{yz}/\sqrt{2} & (R_{xy} + R_{yx}/2) \end{pmatrix} \tag{2.128}$$

For Wigner matrices with $l > 2$, there exist recursion relations that finally reduce to the entries of F and G. The method has some advantages. First, since in our case we do not just need the Wigner D-matrix for a certain l_{max} but also for all l below this value, these are already calculated due to the recursive nature of the algorithm. Second, in simulations the orientational degrees of freedom are often stored as rotation matrices directly. Conversion of those to Euler angles require additional computations and might be inaccurate, especially at "pathological" points (Gimbal lock). It is therefore advantageous that one can use the rotation matrices directly when calculating the Wigner D-matrices. An implementation of this approach is shown in Listing D.1.

Diagonalization method according to Feng et al.

This method was more recently proposed in [Fen+15]. The crucial insight here is the fact that one can rewrite the generator expression that defines the small Wigner d-matrix in a different basis and calculate the corresponding matrix elements. The usual (big) Wigner D-matrix results by adding the two trivial phase factors. Using the spherical basis $|lm\rangle$ with

$$J\,|lm\rangle = l(l+1)\,|lm\rangle \tag{2.129}$$

$$J_z\,|lm\rangle = m\,|lm\rangle \tag{2.130}$$

the small Wigner d-matrix can be written as

$$d^j_{mn}(\theta) = \left\langle jm \middle| e^{-i\theta J_y} \middle| jn \right\rangle \tag{2.131}$$

however, in this basis the operator J_y is not diagonal, so we cannot easily figure out the θ-dependence. Since the eigenstates to the operator J_y, labeled by $|l\mu\rangle_y$, constitute a complete and orthogonal basis we can insert the identity $\sum_\mu |l\mu\rangle_{y\ y}\langle l\mu| = 1$. It follows that

$$d^j_{mn}(\theta) = \left\langle jm \middle| e^{-i\theta J_y} \middle| jn \right\rangle \tag{2.132}$$

$$= \sum_{\mu=-j}^{+j} e^{-i\mu\theta} \langle jm|j\mu\rangle_{y\ y}\langle j\mu|jn\rangle \tag{2.133}$$

The representation of the J_y operator in the spherical basis is given by the matrix

$$J_y = \frac{1}{2i}\begin{pmatrix} 0 & -X_{j+1} & & & \\ X_j & 0 & -X_{-j+2} & & \\ & X_{j-1} & 0 & \ddots & \\ & & \ddots & \ddots & \ddots \\ & & & 0 & -X_j \\ & & & X_{j+1} & 0 \end{pmatrix} \tag{2.134}$$

with $\left\langle jm|J_y|jn\right\rangle = (X_{-n}\delta_{m,n+1} - X_n\delta_{m,n-1})$ and $X_m = \sqrt{(j+m)(j-m+1)}$. Since the matrix is Hermitian we can use standard methods to find the eigenvectors $\{|j\mu\rangle_y\}$, necessary for the evaluation of Equation (2.132). All that remains, is to compute for every matrix entry d^j_{mn} the scalar products $\langle jm|j\mu\rangle_y$ and ${}_y\langle j\mu|jn\rangle$, which are then summed with a weight of $e^{-i\mu\theta}$. Since the scalar products do not depend on the input angle θ we can pre-compute them once at the start of the program and only compute the weighted sums, when the explicit matrix elements are needed. An implementation in the C++ programming language, that makes use of this caching approach, using LAPACK[11] for diagonalization can be found in Listing D.2.

2.6 Symmetries and invariants

In the following we are going to model particles whose interaction potential explicitly depends on their orientation in three-dimensional space. Hence, it is important

[11] LAPACK stands for **L**inear **A**lgebra **Pack**age, which is a standard library for numerical linear algebra, with a focus on efficiency and portability. It is written in Fortran as has bindings for most popular programming languages.

to define first what we mean by a certain orientation and how it is parametrized. In order to specify the coordinates of an object it is necessary to define first a frame of reference relative to which we measure these positions. For the sake of simplicity this can be the usual right-handed Cartesian coordinate system with orthogonal axes x, y and z, represented by xyz. The particle can now be placed into the reference frame, in what we call the standard configuration, because as of now, no further rotation has been applied; the particle is in the unrotated state.

When applying a rotation one has now two choices: Either the reference frame, i.e. the coordinate axes are rotated, or the particle is rotated. In both cases the orientation of the particle relative to the reference frame changes. Performing the rotation in the active sense rotates the particle, while a passive transformation refers to the rotation of the reference frame.

Finally, one has to chose how to parametrize the rotations, in other words, how to realize it by breaking it up into smaller operations. While there are a number of different choices, each with their own advantages and disadvantages we restrict ourselves to the parametrization through Euler angles. Here the complete rotation is decomposed into three individual rotations, each around a certain axis and a certain angle.

Having fixed the rotation we can write in with the shorthand $\Omega = \phi\theta\chi$. Each rotation is then applied from left to right, i.e. $\phi \rightarrow \theta \rightarrow \chi$, and acting on the body-fixed reference frame (borrowing the language from [GG84]). That means, the first rotation around z by ϕ produces a new reference frame with axes $x'y'z'$. The application of θ and χ then produce the axes $x'''y'''z'''$ which is the final configuration.

To repeat, we have

(a) rotate around z with ϕ

(b) rotate around y' with θ

(c) rotate around z'' with χ

The same configuration of the reference frame can however be achieved by another series of rotations. Here, all three sub-rotations are defined relative to the same reference systems (coordinate system), making analysis easier, as one doesn't need to take into account the effect of previous rotations on the axes one is rotating about. The rotations have the same angles, only now is the order of execution reversed, that is $\chi \rightarrow \theta \rightarrow \phi$. That both descriptions are equivalent is not immediately clear but can be shown by performing a transformation using both of the descriptions and comparing the results.

The transformation is now performed as

(a) rotate around z with χ

(b) rotate around y with θ

(c) rotate around z with ϕ

Of great interest to us is the composition of rotations. This comes from the fact that we are interested in the change of certain properties under rotations from the symmetry group of a particle.

In general there are two possibilities for how the combined rotation can be expressed. Let the first rotation be the change from coordinate system $S \rightarrow S'$, described by the Euler angles Ω_1 and the second from $S' \rightarrow S''$, described by the Euler angles Ω_2. The combination of both, i.e. $S \rightarrow S' \rightarrow S''$, can be written as one rotation with the angles Ω. The question is whether the angles Ω_2 are defined with respect to system S or S'.

(a) In the first case where both angles are specified relative to the system S we have for the resulting combination

$$D^l_{kn}(\Omega) = \sum_m D^l_{km}(\Omega_2) D^l_{mn}(\Omega_1) \tag{2.135}$$

(b) In the other case, where Ω_2 is defined relative to S' we get

$$D^l_{kn}(\Omega) = \sum_m D^l_{km}(\Omega_1) D^l_{mn}(\Omega_2) \tag{2.136}$$

For more information the reader can refer to [VMK88], pp. 31–33. We are primarily interested in the first case because we define both the symmetry operation and the individual particle configurations with respect to the same standard configuration.

2.6.1 Invariance condition

Our goal is to expand the angular part of the density distribution $\rho(\mathbf{r}, \Omega)$ in Wigner D-matrices. A symmetry of the density distribution has implications on this expansion. Let in general $f(\Omega)$ be a function that has certain symmetries, such that for a set of angles $\{\Omega_i\}$, which can be viewed as equivalent under the symmetry, the functional values of f are the same. Given the Wigner basis, this function can be

expanded as $f(\Omega) = \sum f^l_{mn} D^l_{mn}(\Omega)$. Because of the symmetry some coefficients are related to each other, therefore it is sufficient to have knowledge of a few coefficients to reconstruct the original function completely. Another way to look at this is to define a new linear combination of Wigner D-matrices, that together transform such that they are like the original function f invariant under the given symmetry operation. Instead of a set of moments that are linearly dependent we arrive at just one "symmetrized" moment, which reduces the bookkeeping-cost.

For the Kern–Frenkel potential, which will be defined later, there are 16 orientational configurations (4 for each particle), that are physically equivalent if we assume the particles' features to be indistinguishable[12]. Starting from the potential, which is a two-body function, we can deduce also the invariant configuration for one-body functions, like the density distribution, by imagining two particles interacting with each other and fixing the configuration of the second particle.

In the following we will construct a new basis of combined Wigner D-matrices that together fulfill any arbitrary point symmetry. We define the new invariant basis as weighted sum over the mth row of the Wigner D-matrix

$$\Delta^l_m(\Omega) = \sum_n d^l_n D^l_{mn}(\Omega)\,, \tag{2.137}$$

with the unknown coefficients d^l_n. For each of the $|G|$ congruent Euler angles the function must evaluate to the same value, as it is invariant by construction. That is

$$\frac{\Delta(\Omega) + \Delta(\Omega') + \cdots}{|G|} = \Delta(\Omega) \tag{2.138}$$

Making use of the transformation relation of Wigner D-matrices we can write

$$\begin{aligned} d^l_n D^l_{mn}(\Omega) &= d^l_n D^l_{mk}(\Omega) \sum_k \frac{1}{|G|} \sum_{R_i \in G} D^l_{kn}(R_i) \\ &= d^l_n \sum_k D^l_{mk}(\Omega) \Pi^{(l)}_{kn} \end{aligned} \tag{2.139}$$

where we introduced the projection operator Π

$$\Pi^{(l)}_{kn} = \frac{1}{|G|} \sum_{R_i \in G} D^l_{kn}(R_i) \tag{2.140}$$

[12] In our case we assume the particles to be "decorated" with patches that are indistinguishable.

By projecting out we arrive at

$$\sum_k \Pi^{(l)}_{nk} d^l_k = d^l_n \tag{2.141}$$

which is an eigenvalue problem with at most $(2l + 1)$ solutions. We are going to label the solutions by an index j as in $d^l_{[j]n}$ which at the same time also labels the different invariant basis functions in every l subspace

$$\Delta^l_{m[j]} = \sum_n d^l_{[j]n} D^l_{mn}(\Omega) \tag{2.142}$$

In order to preserve the normalization of the original Wigner matrices we need to make sure that the following normalization holds

$$\sum_n \left(d^l_{[j]n}\right)^2 = 1 \tag{2.143}$$

To recap, we found those linear combinations of Wigner matrices that transform between each other under rotations *and* are invariant for every rotation from the respective symmetry group. This can be checked by generating a tuple of three Euler angles Ω and calculating the other eleven "congruent" angle tuples Ω', Ω'', ... For every tuple the invariant Wigner matrix gives the same value

$$\Delta^l_{m[j]}(\Omega) = \Delta^l_{m[j]}(\Omega') = \dots \tag{2.144}$$

An alternative perspective can be gained by applying the above process not to a new basis Δ, but to the expansion of a general function on SO(3). Then Equation (2.137) becomes

$$f(\Omega) = \sum_{lmn} f^l_{mn} D^l_{mn}(\Omega) \tag{2.145}$$

with an invariance condition similar to Equation (2.138). The solutions to the resulting linear system are the same $d^l_{[j]n}$ from before, this time with a different interpretation. They tell us the relative magnitude of certain moments with respect to others, as a consequence of the symmetry of the system.

Results for tetrahedral symmetry The method presented above works for any point symmetry, it is only necessary to compute the Euler angles for all rotations in the given group. We are specifically interested in the symmetries generated by the tetrahedral group T. Its group elements are shown in form of permutations

permutation	ϕ	θ	χ	ϕ	θ	χ
(id)	0	0	0	0	0	0
(012)	0	$\pi/2$	$-\pi/2$	$-\pi/4$	$\pi/2$	$-\pi/4$
(031)	π	$\pi/2$	$\pi/2$	$3\pi/4$	$\pi/2$	$3\pi/4$
(132)	π	$\pi/2$	$-\pi/2$	$3\pi/4$	$\pi/2$	$-\pi/4$
(023)	0	$\pi/2$	$\pi/2$	$-\pi/4$	$\pi/2$	$3\pi/4$
(021)	$-\pi/2$	$\pi/2$	π	$-3\pi/4$	$\pi/2$	$-3\pi/4$
(013)	$\pi/2$	$\pi/2$	0	$\pi/4$	$\pi/2$	$\pi/4$
(123)	$-\pi/2$	$\pi/2$	0	$-3\pi/4$	$\pi/2$	$\pi/4$
(032)	$\pi/2$	$\pi/2$	π	$\pi/4$	$\pi/2$	$-3\pi/4$
(02)(13)	0	π	0	$-\pi/2$	π	0
(01)(23)	π	0	0	$-\pi$	0	0
(03)(12)	π	π	0	$\pi/2$	π	0
convention		(A)			(B)	

Table 2.1: Angles for the rotations in the group T for two different conventions of the body frame.

l	basis function	components
3	$\Delta^3_{m[1]}$	$\frac{1}{\sqrt{2}}D^3_{m2} + \frac{1}{\sqrt{2}}D^3_{m\bar{2}}$
4	$\Delta^4_{m[1]}$	$\frac{\sqrt{30}}{12}D^4_{m\bar{4}} + \frac{\sqrt{30}}{12}D^4_{m4} - \frac{\sqrt{21}}{6}D^4_{m0}$
6	$\Delta^6_{m[1]}$	$\frac{\sqrt{7}}{4}D^6_{m\bar{4}} + \frac{\sqrt{7}}{4}D^6_{m4} + \frac{\sqrt{2}}{4}D^6_{m0}$
	$\Delta^6_{m[2]}$	$-\frac{\sqrt{10}}{8}D^6_{m\bar{6}} + \frac{\sqrt{10}}{8}D^6_{m6} - \frac{\sqrt{22}}{8}D^6_{m\bar{2}} + \frac{\sqrt{22}}{8}D^6_{m2}$
7	$\Delta^7_{m[1]}$	$-\frac{\sqrt{39}}{12}D^7_{m\bar{2}} + \frac{\sqrt{33}}{12}D^7_{m\bar{6}} - \frac{\sqrt{39}}{12}D^7_{m2} + \frac{\sqrt{33}}{12}D^7_{m6}$
8	$\Delta^8_{m[1]}$	$-\frac{\sqrt{42}}{24}D^8_{m\bar{4}} + \frac{\sqrt{390}}{48}D^8_{m\bar{8}} + \frac{\sqrt{33}}{8}D^8_{m0} - \frac{\sqrt{42}}{24}D^8_{m4} + \frac{\sqrt{390}}{48}D^8_{m8}$
9	$\Delta^9_{m[1]}$	$\frac{\sqrt{6}}{8}D^9_{m\bar{2}} + \frac{\sqrt{26}}{8}D^9_{m\bar{6}} + \frac{\sqrt{6}}{8}D^9_{m2} + \frac{\sqrt{26}}{8}D^9_{m6}$
	$\Delta^9_{m[2]}$	$-\frac{\sqrt{51}}{12}D^9_{m\bar{4}} - \frac{\sqrt{21}}{12}D^9_{m\bar{8}} + \frac{\sqrt{51}}{12}D^9_{m4} + \frac{\sqrt{21}}{12}D^9_{m8}$

Table 2.2: Linear combinations of Wigner matrices that are invariant under all tetrahedral group transformations up to $l = 9$.

l	non-zero D^l_{mn}	independent $\Delta^l_{m[j]}$
0	0	0
1	0	0
2	0	0
3	2	1
4	3	1
5	0	0
6	7	2
7	4	1
8	5	1
9	8	2

Table 2.3: The number of non-vanishing D^l_{mn} and linearly independent $\Delta^l_{m[j]}$ up to $l = 9$, resulting from tetrahedral symmetry, for a fixed m. For a general inhomogeneous situation, $m = -l...l$, whereas in the flat wall geometry (cylindrical symmetry), $m = 0$.

in Table 2.1. The group is isomorphic to the group $A(4)$, the alternating group of degree four which consists of all even permutations on four elements, shown in the first column of the table. This can be seen by identifying the four vertices each with one number from the set $\{0, 1, 2, 3\}$. A symmetry rotation of the tetrahedron then corresponds to a permutation of the vertices. As we only consider proper rotations, this corresponds to the alternating group and not to the symmetric group $S(4)$, which contains 24 elements in total. A graphical representation of this isomorphism is shown in Figure 2.3.

The first couple of symmetric basis functions are shown in Table 2.2, with the corresponding number of linearly independent basis functions in Table 2.3. Due to choice of patch positions, all matrix elements appear symmetrically, i.e. D^l_{mn} and $D^l_{m\underline{n}}$ appear with the same prefactor. We can see from the tables that there are no basis functions for $l < 3$ as these would not transform in the appropriate way (for $l = 1$ for example, they would transform as a vector).

2.6.2 Group theoretical description

Alternatively, one can approach the problem from a group theoretical point of view. For an introduction to group theory in the context of physics, one can choose from

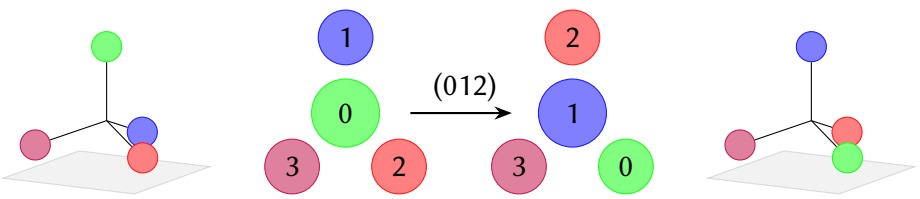


Figure 2.3: Action of the permutation on the positions of the patches. The diagrams on the l.h.s. show the particle from a top view. A three-dimensional representation is also given for easier visualization.

a number of texts, e.g. the classic reference [Ham89] or for a more recent treatment [RD19]. For the sake of brevity, we will not go into too much detail regarding definitions and proofs, but only offer a condensed overview of the most important concepts.

A group G is a set of elements g_i, and group multiplication that is closed

$$g_1 g_2 \mapsto g_3 \in G \tag{2.146}$$

Further, G needs to contain the neutral element e, such that $\forall g \in G$: $eg = g$ and the inverse g^{-1} for every element $g \in G$, such that $g^{-1}g = e$. Then, the group's structure (i.e. the way the elements behave when multiplied, summarized in a group table) can be realized by linear operators (matrices) acting on some vector space V, given by the map $\rho : G \to \mathrm{GL}(V)$. For every group there is a large number of families of matrices that follow the group structure. However, one can single out so-called irreducible representations, which have the property that there is no smaller subspace that is G-invariant, other than the zero vector space.

Given any general representation, it is possible to decompose it into irreducible representations (irreps). This mean that the irreducible representations are the building blocks for all the representations of a given group. Specifically, we are interested in the decomposition of reducible representations into irreducible representations.

As mentioned earlier, the irrep of SO(3) is given by the Wigner D-matrices. In contrast to the well-known point groups, there are infinitely many conjugacy classes[13], one for every rotation angle $\psi \in [0, 2\pi)$. Note, that this rotation angle ψ does not correspond to any of the Euler angles but equals the rotation angle in the angle–axis parametrization[14].

[13] Class of equivalent group elements.

[14] This parametrization is given a rotation axis $\mathbf{a}$ that can point in any direction and a rotation angle ψ.

One can further define the character of a group element by the function χ, given by

$$\chi_\rho(g) = \mathrm{Tr}(\rho(g)) \tag{2.147}$$

where Tr is the trace operation defined on linear operators. Due to the properties of the trace, it is invariant under basis changes of V. For a given group, the character table lists the character of all conjugacy classes and irreps. The character of the identity elements (denoted by Id) is equal to the dimension of the representation.

The characters of the Wigner D-matrix representation are given by [VMK88]

$$\chi^j(\psi) \equiv \sum_m D^j_{mm}(0,\psi,0) = \sum_m d^j_{mm}(\psi) = \frac{\sin\left((2j+1)\psi/2\right)}{\sin(\psi/2)} \tag{2.148}$$

The character tables for most physically relevant point groups can be found in standard text books like [Ham89]. For the tetrahedral group T, the table is given by Table 2.4. We see that the tetrahedral group has three irreducible representations in total, usually denoted by A, E and T. In the trivial representation A, every element is given by the real number 1, it is therefore one-dimensional, and all characters are also equal to one. The representations E and T and respectively two- and three-dimensional[15], though E can be seen as combination of two one-dimensional representations, where one is the complex conjugate of the other. In the following, we will only be interested in the trivial representation A. The formula for obtaining

	Id	$4\,C_3$	$4\,C_3^2$	$3\,C_2$
A	1	1	1	1
E	1	θ	θ^*	1
	1	θ^*	θ	1
T	3	0	0	-1

Table 2.4: Character table of the tetrahedral group, showing the character of every group element of one of the three irreducible representations A, E and T. The parameter θ corresponds to the value $e^{i2\pi/3}$.

the multiplicity m^k, which tells us how often an irreducible representation appears in a given reducible representation is given by

$$m^k = \frac{1}{|G|}\sum_{g\in G}\chi^\Gamma(g)\overline{\chi^{\Gamma^k}(g)} \tag{2.149}$$

[15]Note, that the dimensionality of an irrep can be read off from the group table Table 2.4. The element Id is always the identity matrix and the trace therefore is the dimensionality of the representation.

The coefficients for the characters of the irrep χ^{Γ^k} can be found in Table 2.4. The ones for the reducible representation (χ^{Γ}) can be calculated using Equation (2.148) and the angles for the group elements given in Table 2.4. The correspondence between group element and rotation angle is given by

- the identity, corresponding to 0°
- 4 C_3, corresponding to 120°
- 4 C_3^2, corresponding to, again 120°
- and 3 C_2, corresponding to 180°

j	Id	8 $C_3^{(2)}$	3 C_2
1	1	1	1
1	3	0	-1
2	5	-1	1
3	7	1	-1
4	9	0	1
5	11	-1	-1
6	13	1	1
7	15	0	-1
8	17	-1	1
9	19	1	-1

Table 2.5: Characters χ^j of the tetrahedral group in the Wigner D-matrix representation.

As an example, let us calculate m^A for the case $j = 3$, in other words how often the irrep A of the tetrahedral group appears in the D-matrix representation with $j = 3$. As the characters for the irrep A are all equal to one, i.e. $\chi^{\Gamma^k} = 1$, the formula reduces to a sum of χ^{Γ} only

$$m^A = \frac{1}{12}\left[7 \cdot 1 + 8 \cdot 1 + 3 \cdot (-1)\right] = 1 \tag{2.150}$$

meaning it appears once. The only other irrep that appears is T with $m^T = 2$. As a result, the decomposition for $j = 3$ is given by

$$A \oplus 2T \tag{2.151}$$

The decomposition for other values of j up to $j = 9$ are shown in Table 2.6. This means that we can find the same information that is contained in Table 2.3, in a

j	decomposition
0	**A**
1	T
2	E ⊕ T
3	**A** ⊕ 2 T
4	**A** ⊕ E ⊕ 2 T
5	E ⊕ 3 T
6	2 **A** ⊕ E ⊕ 3 T
7	**A** ⊕ E ⊕ 4T
8	**A** ⊕ 2 E ⊕ 4 T
9	2 **A** ⊕ E ⊕ 5T

Table 2.6: Decomposition of the individual representations for every j up to the value $j = 9$. The trivial representation A is highlighted in bold.

much faster way, by computing how often the representation A appears in the decomposition belonging to the number j, i.e. the number in front of A in Table 2.6 is equal to the entries in Table 2.3.

The projector onto the irrep k from the reducible representation Γ is given by

$$P_\Gamma^k = \frac{\ell_{\Gamma^k}}{|G|} \sum_{g \in G} \overline{\chi^{\Gamma^k}(g)} \Gamma(g) \tag{2.152}$$

We can see that if we chose as irrep the trivial representation A (where all $\chi^{\Gamma^k}(g)$ are equal to one), then the projector corresponds exactly to the one presented earlier in Equation (2.140).

2.6.3 Rotational invariants

Similar symmetry considerations apply also when working with two-body functions like the radial distribution function, which now depend on the relative distance $\mathbf{r}$ between the particles and their respective orientations Ω_i

$$g(\mathbf{r}, \Omega_1, \Omega_2) \quad \text{or} \quad g(r, \Omega_1, \Omega_2, \hat{\mathbf{r}}) \tag{2.153}$$

Again, there is the possibility of expanding the orientation degrees of freedom in Wigner D-matrices, however there are multiple possibilities. Instead of expanding

g in terms like[16]

$$D^l_{mn}(\Omega_1)D^u_{vw}(\Omega_2)D^a_{b0}(\hat{\mathbf{r}}) \tag{2.154}$$

it makes sense to define certain linear combinations of Wigner D-matrices that are invariant under rotation of the whole two-body system. The reason for this is that "genuine" two-body functions like the RDF have exactly this property, since they are independent of the reference system that one uses. Expanding in such an invariant basis automatically leads to expansion coefficients that are the same in every rotated coordinate system, which has two advantages. First, it is not necessary to specify a special frame, as they are now all equivalent, and second, if we need to evaluate a certain quantity we can choose a frame that makes the computation particularly easy. This type of invariant expansion was first introduced in [BT72] and then later worked out in more detail in [Blu72; Blu73; BT88].

The invariant expansion of a pair-function f is given

$$f(\mathbf{r}, \Omega_1, \Omega_2) = \sum_{mnl\mu\nu} f^{mnl}_{\mu\nu}(r_{12})\Phi^{mnl}_{\mu\nu}(\Omega_1, \Omega_2, \hat{\mathbf{r}}) \tag{2.155}$$

where the invariant basis functions Φ are given by

$$\Phi^{mnl}_{\mu\nu}(\Omega_1, \Omega_2, \hat{\mathbf{r}}) = c_m c_n \sum_{\mu'\nu'\lambda'} \begin{pmatrix} m & n & l \\ \mu' & \nu' & \lambda' \end{pmatrix} \overline{D^m_{\mu'\mu}}(\Omega_1)\overline{D^n_{\nu'\nu}}(\Omega_2)\overline{D^l_{\lambda' 0}}(\hat{\mathbf{r}}) \tag{2.156}$$

with the normalization[17] $c_m = \sqrt{2m+1}$. The interpretation is that the $3j$-symbols take care of summing the right combinations of Wigner D-functions such that taken together any extra coefficients resulting from a rotation of the reference frame cancel, and the result becomes invariant. This is in fact a general property. For spherical tensors [i.e. they transform according to Equation (2.115)] $A_{m\mu}, B_{n\nu}, C_{l\lambda}$ the product

$$\sum_{\mu\nu\lambda} \begin{pmatrix} m & n & l \\ \mu & \nu & \lambda \end{pmatrix} A_{m\mu}B_{n\nu}C_{l\lambda} \tag{2.157}$$

is always a rotationally invariant scalar, as shown in Appendix A of [GG84]. We can therefore take advantage of the fact that the Wigner D-matrices are spherical tensors themselves. The connection to the Clebsch–Gordan coefficients can be made by the relation

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \equiv \frac{(-1)^{j_1-j_2-m_3}}{\sqrt{2j_3+1}} \langle j_1 m_1 j_2 m_2 | j_3(-m_3)\rangle \tag{2.158}$$

[16] The radial vector $\hat{\mathbf{r}}$ in the argument of the D-matrix should be interpreted as the two Euler angles corresponding to the direction of the radial vector.

[17] Note that [GG84] uses a different normalization. We prefer the one given as then the moment $g^{000}_{00} = g(r)$, without further factors. It is also used in [Bel17].

however, it may be more practical to work with the $3j$-symbols as their symmetries are easier to work with. For efficient numerical evaluation of the $3j$-symbols see [JF16].

The coefficients $f_{\mu\nu}^{mnl}$ are found by projecting the full function f onto the basis functions

$$f_{\mu\nu}^{mnl} = (2l+1) \int \mathrm{d}\Omega_1 \, \mathrm{d}\Omega_2 \, \mathrm{d}\hat{\mathbf{r}} f_{\mu\nu}^{mnl}(\hat{\mathbf{r}}, \Omega_1, \Omega_2) \overline{\Phi_{\mu\nu}^{mnl}}(\Omega_1, \Omega_2, \hat{\mathbf{r}}) \tag{2.159}$$

From this viewpoint it is straightforward to include the tetrahedral one-body symmetry that we treated before. For that, we replace every occurrence of normal Wigner matrix with that of the invariant one, $D \to \Delta$

$$\Phi_{[ij]}^{mnl}(\Omega_1, \Omega_2, \hat{\mathbf{r}}) = c_m c_n \sum_{\mu'\nu'\lambda'} \begin{pmatrix} m & n & l \\ \mu' & \nu' & \lambda' \end{pmatrix} \overline{\Delta_{\mu'[i]}^{m}}(\Omega_1) \overline{\Delta_{\nu'[j]}^{n}}(\Omega_2) \overline{D_{\lambda' 0}^{l}}(\hat{\mathbf{r}}) \tag{2.160}$$

Since the basis functions Δ are labeled by an index the same carries over to Φ, replacing the previous indices $\mu\nu$.

While the expansion coefficients of the function f are independent of rotations of the lab frame, they do depend on the choice of the "unrotated" frame for the individual particles. This means that different conventions lead in general to different expansion coefficients; they are however linearly dependent. In our case this choice would translate to the positioning of the patches on the individual particle.

2.6.4 Intermolecular frame

In order to calculate anisotropic pair quantities like the radial distribution function we are effectively creating histograms of functions which depend on eight variables, consisting of twice the three Euler angles plus the two (spherical) coordinates of the connection vector between them.

For ease of reading we repeat here the invariant expansion of a function f

$$f(\hat{\mathbf{r}}, \Omega_1, \Omega_2) = \sum_{mnl\mu\nu} f_{\mu\nu}^{mnl}(r_{12}) \Phi_{\mu\nu}^{mnl}(\Omega_1, \Omega_2, \hat{\mathbf{r}}_{12}) \tag{2.161}$$

We can now take advantage of the fact that the basis functions $\Phi_{\mu\nu}^{mnl}$ are the same in every reference frame. If we found a frame that makes the calculation of the coefficients particularly easy we can transform into that frame, perform the calculation

and then transform back into the original frame. The basis functions Φ assume a particularly simple form when the separation vector $\hat{\mathbf{r}}$ points into the z-direction, as then

$$D^l_{mn}(0,0,0) = \delta_{mn} \quad \text{or} \quad D^l_{\lambda' 0}(0,0,0) = \delta_{\lambda' 0} \tag{2.162}$$

We call this reference frame the intermolecular (IM) frame and label the Euler angles with respect to this frame with a lowercase ω, i.e. the rotation of particle i is given by Ω_i in the lab frame and ω_i in the intermolecular frame. The transformation from $\Omega \longrightarrow \omega$ depends on the direction $\hat{\mathbf{r}}$ and will be given later. Note that numerically

$$\Phi^{mnl}_{\mu\nu}(\omega_1, \omega_2, 0) = \Phi^{mnl}_{\mu\nu}(\Omega_1, \Omega_2, \hat{\mathbf{r}}) \tag{2.163}$$

In the IM frame the basis functions reduce to

$$\Phi^{mnl}_{\mu\nu}(\omega_1, \omega_2, 0) = c_m c_n \sum_{\chi} \begin{pmatrix} m & n & l \\ \chi & \underline{\chi} & 0 \end{pmatrix} \overline{D^m_{\chi\mu}}(\omega_1) \overline{D^n_{\underline{\chi}\nu}}(\omega_2) \tag{2.164}$$

with the expansion becoming

$$f(\omega_1, \omega_2, r) = \sum_{mnl\mu\nu} f^{mnl}_{\mu\nu}(r) \sum_{\chi} \begin{pmatrix} m & n & l \\ \chi & \underline{\chi} & 0 \end{pmatrix} \overline{D^m_{\chi\mu}}(\omega_1) \overline{D^n_{\underline{\chi}\nu}}(\omega_2) \tag{2.165}$$

We introduce a new collection of expansion coefficients with the relation

$$f^{mn}_{\mu\nu\chi}(r) = \sum_{l} \begin{pmatrix} m & n & l \\ \chi & \underline{\chi} & 0 \end{pmatrix} f^{mnl}_{\mu\nu}(r) \tag{2.166}$$

which goes under the name of χ-transform [BT72]. This relation can be inverted, yielding

$$f^{mnl}_{\mu\nu}(r) = (2l+1) \sum_{l} \begin{pmatrix} m & n & l \\ \chi & \underline{\chi} & 0 \end{pmatrix} f^{mn}_{\mu\nu\chi}(r) \tag{2.167}$$

Rewriting Equation (2.165) in the χ-transformed coefficients yields

$$f(\omega_1, \omega_2, r) = \sum_{mn\mu\nu\chi} f^{mn}_{\mu\nu\chi}(r) \overline{D^m_{\chi\mu}}(\omega_1) \overline{D^n_{\underline{\chi}\nu}}(\omega_2) \tag{2.168}$$

which means that we can find $f^{mn}_{\mu\nu\chi}(r)$ by the following projection

$$f^{mn}_{\mu\nu\chi}(r) = \int \mathrm{d}\omega_1 \, \mathrm{d}\omega_2 f(\omega_1, \omega_2, r) D^m_{\chi\mu}(\omega_1) D^n_{\underline{\chi}\nu}(\omega_2) \tag{2.169}$$

which is an integral over only two triplets of Euler angles. By further rotating the IM frame such that ϕ_1 is equal to 0, we can reduce the integration to just five dimensions. The original quantity of interest can then be computed by the inverse χ-transform Equation (2.167).

Performing the transformation into the IM-frame Our goal is to transform the orientations in the lab frame Ω_i to orientations in the IM frame, where the connection vector points in the z-direction. The representation for the orientations used in our simulation code is given by rotation matrices. Therefore, we are going to restrict ourselves to the transformation of the rotation matrices and not the Euler angles in general. Anyhow, conversion between the two is always possible.

Assume that the vector pointing from particle A to particle B is given in the lab system (defined by the simulation box) by the vector $\mathbf{r}_L$ We want to find the rotation matrix $R_{L\to\mathrm{IM}}$ that makes $\mathbf{r}_L$ parallel to the z-direction, thus defining the intermolecular frame. It is clear that $R_{\mathrm{IM}\to L} = R^{-1}_{L\to\mathrm{IM}}$ and that

$$R_{\mathrm{IM}\to L}\mathbf{e}_z = \mathbf{r}_L \tag{2.170}$$

Given the Cartesian coordinates of $\mathbf{r}_L$ it is easy to find the corresponding spherical coordinates ϕ, θ for which

$$R_z(\phi)R_y(\theta)\mathbf{e}_z = \mathbf{r}_L \tag{2.171}$$

By $R_i^{-1}(\alpha) = R_i(-\alpha)$ and orthogonality of rotation matrices we have

$$R_{L\to\mathrm{IM}} = R_y(-\theta)R_z(-\phi) \tag{2.172}$$

The patches of any particle in the lab frame result from multiplication of the orientation matrix of said particle with the patch positions in the unrotated frame $\mathbf{p}^i$

$$\mathbf{p}^i_L = R(\Omega)\mathbf{p}^i \tag{2.173}$$

The orientation matrix in the IM frame can therefore be written as

$$R_{L\to\mathrm{IM}}R(\Omega) = R_y(-\theta)R_z(-\phi)R(\Omega) = R(\omega) \tag{2.174}$$

3 Kern–Frenkel (KF) model and patchy particles

3.1 The Kern–Frenkel model

The Kern–Frenkel potential is a simple, yet rich model for associating particles. Due to its rather easy implementation and flexibility is often used as a model for investigations of patchy particles. Control over the bonding geometry further allows for comparison with results from Wertheim's thermodynamic perturbation theory, that will be discussed later. The KF potential consists of a hard-sphere potential plus an anisotropic attractive part that is due to patches attached to the surface of the sphere. The patches are described as cones, which have their tip in the particle center and extend into the radial direction, their base surface being curved. The number of cones is variable, here we restrict ourselves to four patches arranged in tetrahedral configuration.

The Kern–Frenkel potential is defined by

$$V^{\mathrm{KF}}(\mathbf{r}_{12}, \Omega_1, \Omega_2) = \phi_{\mathrm{hs}}(r_{12}) + \phi_{\mathrm{sw}}(r_{12}) \sum_{\alpha,\beta=1}^{N_{\mathrm{p}}} \phi_{\mathrm{p}}(\mathbf{r}_{12}, \hat{\mathbf{r}}_1^{\alpha}(\Omega_1), \hat{\mathbf{r}}_2^{\beta}(\Omega_2)) \tag{3.1}$$

showing that the radial and angular degrees of freedom factorize, making the model easier to treat analytically. Here, N_{p} is the number of patches and $\phi_{\mathrm{sw}}(r)$ is a square well potential

$$\phi_{\mathrm{sw}}(r) = \begin{cases} -\epsilon & \text{if } r \in [\sigma, \sigma + \delta] \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

with the attractive range δ and attraction depth ϵ. ϕ_{p} is the orientation-dependent patch–patch part

$$\phi_{\mathrm{p}}(\mathbf{r}_{12}, \hat{\mathbf{r}}_1^{\alpha}(\Omega_1), \hat{\mathbf{r}}_2^{\beta}(\Omega_2)) = \begin{cases} 1 & \text{if both } \begin{cases} \hat{\mathbf{r}}_{12} \cdot \hat{\mathbf{r}}_1^{\alpha} > \cos\theta_{\max} \\ -\hat{\mathbf{r}}_{12} \cdot \hat{\mathbf{r}}_2^{\beta} > \cos\theta_{\max} \end{cases} \\ 0 & \text{else} \end{cases} \tag{3.3}$$

characterized by the opening angle $\theta_{\max}$ of the "cones" that interact attractively between the particles[1]. For the directions of the patches we chose the following convention

$$P = \frac{1}{\sqrt{3}} \begin{pmatrix} -\sqrt{2} & \sqrt{2} & 0 & 0 \\ 0 & 0 & -\sqrt{2} & \sqrt{2} \\ 1 & 1 & -1 & -1 \end{pmatrix}. \tag{3.4}$$

where each column is the vector pointing from the particle center into the respective patch direction.

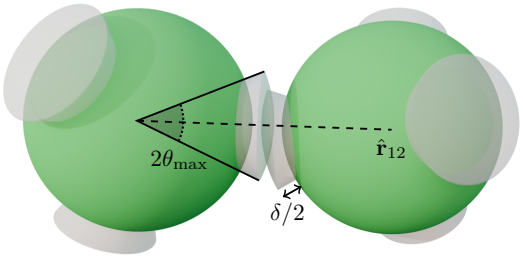


Figure 3.1: Three-dimensional render of two Kern–Frenkel particles being in the bonded state. The vector $\hat{\mathbf{r}}_{12}$ is the connection from one center to the other and the parameters $\theta_{\max}$ and δ are parameters of the KF potential.

Single bond condition Our plan is to use Wertheim's TPT1 as a reference for further developments of the excess free energy. One of the principal assumptions of TPT1 is that only one bond is allowed per patch. Because the KF potential allows in principle parameters for which *one* patch of particle A is bonded *both* to a particle B and a separate particle C, we need to tune the cone opening angle and the attraction range jointly. The single bond constraint is fulfilled when the condition

$$\sin(\theta_{\max}) \leq [2(1+\delta)]^{-1} \tag{3.5}$$

holds, see [RRR18]. We chose the parameters as given by

$$\delta = 0.119 \tag{3.6}$$

$$\cos\theta_{\max} = 0.895 \tag{3.7}$$

For a true to scale representation of these parameters, see Figure 3.1. The critical temperature T_c, corresponding to this choice of parameters, can be estimated from

[1] In order to evaluate the bonding part in simulations only two cosine evaluations are necessary for every patch pair.

simulations [FS07] and is equal to $T_c = 0.1682 \pm 0.0005$, $\eta_c = 0.139 \pm 0.004$. For comparison, the critical values for TPT1 are given by [Sto19]

$$T_c = 0.161\,, \quad \eta_c = 0.172 \tag{3.8}$$

The surface coverage χ is the percentage of particle surface that is covered by the patches, and depends on the cone angle and the number of patches n. It is given by

$$\chi = n \sin^2\left(\frac{\theta_{\max}}{2}\right) = \frac{n}{2}\left(1 - \cos(\theta_{\max})\right) \tag{3.9}$$

The bonding volume v_b, defined as the part of the rotational phase space, for which two adjacent particles are in the bonded state, is given by

$$v_b = \chi^2 \frac{4\pi\sigma^3}{3}\left[(1+\delta)^3 - 1\right] \tag{3.10}$$

Note, that the surface coverage can also be written as multiple of the zero-moment of the Kern–Frenkel potential $V \equiv V^{\mathrm{KF}}$, i.e.

$$\chi^2 \int \mathrm{d}\Omega_1\, \mathrm{d}\Omega_2\, \mathrm{d}\hat{\mathbf{r}} = \int \mathrm{d}\Omega_1\, \mathrm{d}\Omega_2\, \mathrm{d}\hat{\mathbf{r}} V(\Omega_1, \Omega_2, \hat{\mathbf{r}}) = V_{00}^{000} \tag{3.11}$$

$$= 256\pi^2\chi^2 \tag{3.12}$$

More information on the moments of the KF potentials will be presented in Section 4.4.1.

3.2 Physics of patchy particles

In recent years patchy particles have attracted much attention due to their interesting properties. Especially particles with four patches arranged in a tetrahedral configuration represent a valuable test system when trying to understand real anisotropic fluids, such as water. While being a relatively simple model, patchy particles display a rich physical behavior, for an overview see [Rus+22].

The important difference to other often used models for fluids, say interaction through square wells or electrostatic forces, is the fact that the interaction is orientation dependent and not just radially symmetric. This modification makes it possible for the fluid to be in a percolated state, with system-spanning particle clusters and relatively long lifetimes.

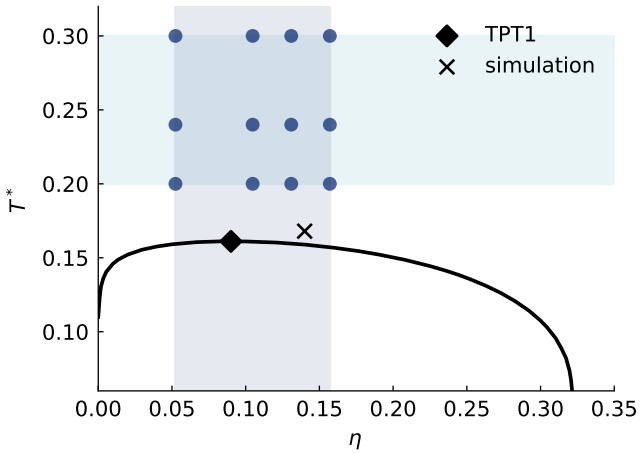


Figure 3.2: Phase diagram of TPT1 together with the critical point of MC simulations for $M = 4$ patches and the parameters mentioned in the main text. The binodal of the KF systems is qualitatively similar to the result of TPT1, though slightly wider and flatter, see [Bia+08]. The blue dots represent the state points, at which our training data is collected in the following treatment.

Empty liquids In order to understand the novel behavior of patchy particles, let us first recapitulate the phase behavior of a liquid, that interacts through an isotropic potential, e.g. a square well or the Lennard-Jones potential. Below the critical temperature the liquid can coexist at two different phases, the low-density gas phase (ρ_g) and the higher density liquid phase (ρ_l). The liquid branch of the binodal is rather steep, and densities grow rapidly as the temperature decreases, making the coexistence region rather wide. If we consider a point above the critical temperature and a density larger than the critical density, it follows that the liquid becomes unstable (by entering the coexistence region) when lowering the temperature. It is therefore not possible to reach the state of a low-density, stable liquid.

The situation changes in the case of patchy particles. Here, the binodal decreases in a much steeper fashion, and therefore does not extend that far out into the high density region, see Figures 3.2 and 3.3. The exact width depends on the number of patches (the valence), with fewer patches having a narrower binodal and lower critical temperature. The much smaller width of the binodal opens up the possibility of forming *equilibrium gels* in the region right of the binodal but still before the solid region. Compared to packing fractions of liquids interacting isotropically, the packing fractions here are much lower, hence the term “empty”. The reason for this

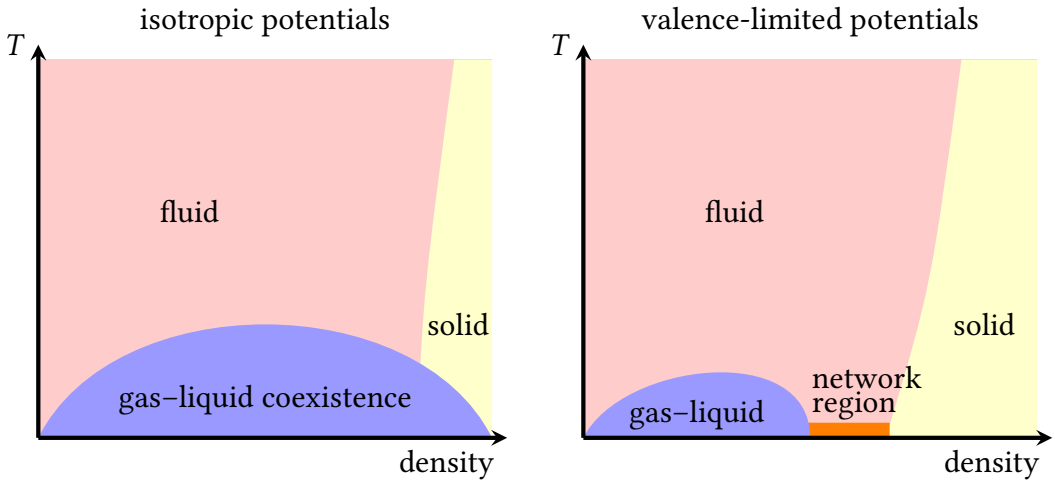


Figure 3.3: Sketches of phase diagrams of liquids interacting through an isotropic potential and those interacting through valence-limited (VL) potentials like the Kern–Frenkel potential. Notice, how in the isotropic case the binodal extends into the glass region, while in the VL-case, it ends before it, with a region of equilibrium gels forming in between. The plot omits the solid phase that is located at very high densities. Figure adapted from [Sci08].

behavior is the limited valence of the particles. In contrast to isotropic particles, increasing the number of potential bonding partners would not result in a lower energy state, as all bonding sites are already saturated. Interesting behavior can be observed when mixing two- and three-patch particles, such that the mean number of patches $\langle M \rangle$ varies in the interval $2 < \langle M \rangle < 3$. Towards the limit $\langle M \rangle \to 2$ the liquid density approaches that of the vapor, meaning that by tuning the number of patches one can reach in principle arbitrarily low density fluid phases [Bia+06; HTG11].

3.3 Monte Carlo simulations

The "ground truth" data is generated using computer simulations of the system. In principle, we have either the choice of doing molecular dynamics simulation (integration of the equation of motion), or Monte Carlo simulations, which is a method to generate samples of phase space coordinates of the system, that follow the equilibrium probability distribution of the system at a certain temperature. Because

the interparticle potential is discontinuous, which leads to diverging forces, it is simpler to simulate ensembles of patchy particles using Monte Carlo techniques.

The basic working principle of a (canonical) MC simulation is a follows. The physical state of the N particles making up the system is initialized, for example, by random uniform sampling over the possible coordinates. Let this state be called $\mathbf{s}_0$. Depending on the specific method that is used, a new state is proposed, e.g. by displacing and rotating a particle by a small amount (also called a rototrans move). Afterwards the ensemble energy corresponding to this new state $E(\mathbf{s}_{i+1})$ is calculated. Depending on the energy difference $\Delta E = E(\mathbf{s}_{i+1}) - E(\mathbf{s}_i)$ the new state is accepted. This is always the case if $\Delta E \leq 0$ (energetically more favorable) and with probability

$$p_{i+1} = e^{-\beta \Delta E} \tag{3.13}$$

if the total energy increases, i.e. $\Delta E > 0$. The bigger the change, the less likely it is that the move is accepted. This is further amplified when the temperature is low (large β). For this reason it can become difficult to sample effectively from systems at lower temperature because simple rototrans moves are usually rejected.

The correct statistical sampling is implemented by first drawing a uniform number u between 0 and 1, and accepting the move if $u \leq p_{i+1}$; otherwise the old state is kept. After a sufficient number of moves the systems is thermalized and represents a sample of the system at equilibrium. Subsequent samples also follow the equilibrium distribution, they are however highly correlated, because there is little change between individual snapshots. In order to get rid of these unphysical correlations a number of moves is discarded after every sample and not considered when calculating expectation values. These states are only used to advance the Markov chain and decrease the correlation. For more detailed information on Monte Carlo techniques, specifically in the context of liquid state theory we refer to the excellent textbooks [AT17; FS23].

For low density systems it becomes difficult to reach equilibrium states using only rotation and translations (rototrans) moves. There exist more sophisticated MC moves, e.g. the AVB move, that transports particles into the bonding region of other particles. In order to conserve detailed balance, the reverse move of breaking up clusters has to happen with the same probability. We use the publicly available code called PatchyParticles by Rovigatti et al. [RRR18]. It supports multiple ensembles and MC moves, especially also the AVB move.

Random orientations During the implementation of simulations of anisotropic particles one needs to decide on how to computationally deal with the rotational

degrees of freedom. Saving the individual Euler angles for every particle reduces memory consumption but makes it necessary to compute the corresponding rotation matrix to evaluate the patch orientation. Therefore, it is more sensible to store either the rotation matrix itself or alternatively, use quaternions to parametrize the rotations. While the latter is often preferred for numerical stability and easy random sampling, it adds a layer of indirection. The PatchyParticles code therefore uses the rotation matrix representation internally. One can save some memory by only keeping account of two rows and then compute the third when necessary as the rotation matrix is orthogonal. Due to inexact floating point arithmetic it is necessary to re-orthogonalize the rotation matrices after a number of moves. One difficulty that arises from this choice of parametrization is the process of sampling a new rotation that is close to the previous one. One might implement this move by sampling a random direction on the sphere $\mathbf{n}$ and then rotate this around an angle ϕ, leading to the matrix $R(\mathbf{n}, \phi)$, which is then multiplied by the old rotation matrix. It is important to realize though, that the angle ϕ is in this case not uniformly distributed in $[0, 2\pi]$, i.e. $\phi \nsim \mathcal{U}(0, 2\pi)$, but following the probability density $\sin(\frac{1}{2}\phi)^2$. For a proof see [Mil65].

An alternative way is to sample $\phi, \chi \sim \mathcal{U}(0, 2\pi)$, $z \sim \mathcal{U}(-1, 1)$ and setting $\theta = \arccos(z)$, as described in [Kuf04]. More information about random sampling of orientations using different parametrizations can be found in Ref. [Yer+10].

3.4 Orientational structure of the KF liquid

Due to the anisotropic interaction potential, it is clear the system distribution functions will also become anisotropic. We specifically look at the radial distribution function $g(r)$ and the one-particle density distribution ρ at a wall and around a (hard/square well) radial tracer particle.

3.4.1 Homogeneous case

In the case of a homogenous liquid, i.e. without an external potential, we are interested in the fully resolved (molecular) pair distribution function $g(\mathbf{r}_1, \mathbf{r}_2, \Omega_1, \Omega_2) = \rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2, \Omega_1, \Omega_2)/\rho^2$. With the molecular pair density defined by

$$\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2, \Omega_1, \Omega_2) = \left\langle \sum_{i=1}^{N} \sum_{j \neq i}^{N} \delta(\mathbf{r}_1 - \mathbf{r}_i)\delta(\mathbf{r}_2 - \mathbf{r}_j)\delta(\Omega_1 - \Omega_i)\delta(\Omega_2 - \Omega_j) \right\rangle \quad (3.14)$$

Since the systems is homogeneous, the RDF does not depend on the two individual positions $\mathbf{r}_1, \mathbf{r}_2$, but only the separation vector $\mathbf{r}$

$$\begin{aligned}\rho^{(2)}(\mathbf{r}, \Omega_1, \Omega_2) &= \frac{1}{V}\int \mathrm{d}\mathbf{r}' \rho^{(2)}(\mathbf{r}' + \mathbf{r}, \mathbf{r}', \Omega_1, \Omega_2) \\ &= \frac{1}{V}\left\langle \sum_{i=1}^{N}\sum_{j\neq i}^{N} \delta(\mathbf{r} - \mathbf{r}_{ij})\delta(\Omega_1 - \Omega_i)\delta(\Omega_2 - \Omega_j)\right\rangle\end{aligned} \tag{3.15}$$

The dependence on the individual orientations Ω_i remains, however. The pair distribution function can then be written as

$$\begin{aligned}g(\mathbf{r}_{12}, \Omega_1, \Omega_2) &= \frac{\rho^{(2)}(\mathbf{r}_{12}, \Omega_1, \Omega_2)}{\rho^2} \\ &= \frac{V}{N(N-1)}\left\langle \sum_{i=1}^{N}\sum_{j\neq i}^{N} \delta(r - r_{ij})\delta(\Omega_1 - \Omega_i)\delta(\Omega_2 - \Omega_j)\right\rangle \frac{1}{4\pi r^2}\end{aligned} \tag{3.16}$$

In the usual case of isotropic particles one can proceed by binning simulation data into radial bins centered at $r + \delta r$. Then

$$g(r_k) = n(r_k)/n^{\mathrm{id}}(r_k) \tag{3.17}$$

where n^{id} is the number of particles in the (finite) radial shell at radius r_k, calculated by

$$n^{\mathrm{id}}(r_k) = \frac{4\pi\rho}{3}\left[(r + \delta r)^3 - r^3\right] = V_k\rho \tag{3.18}$$

Note that in the case of infinitesimal δr we reproduce the formula of Equation (3.16) with the identification $V_k \to 4\pi r^2$, $n(r_k) \to \frac{\langle\cdots\rangle}{N}$ and $N/V \to \rho$. In the case of anisotropic particles, binning for the six Euler angles would turn out to be rather expensive and statistically unfavorable. We therefore expand the rotational degrees of freedom in the rotational invariants defined earlier. This way we return to one-dimensional histograms that only depend on the radial distance r. The expansion reads

$$g^{lmn}_{\mu\nu}(r, \Omega_1, \Omega_2) = \int \mathrm{d}\Omega_1\, \mathrm{d}\Omega_2\, \mathrm{d}\hat{\mathbf{r}} g(\mathbf{r}, \Omega_1, \Omega_2)\overline{\Phi^{lmn}_{\mu\nu}}(\Omega_1, \Omega_2, \hat{\mathbf{r}}) \tag{3.19}$$

In order to save computing time, we make use of the intermolecular frame representation introduced earlier. For every pair of particles with orientations Ω_1, Ω_2 and separation vector $\mathbf{r}_{12}$ we first find the rotation that is necessary to bring $\mathbf{r}_{12}$ into the z-direction. The rotation is determined by the procedure described in Equation (2.174). For the moments of g in the intermolecular frame it follows that

$$g^{mn}_{\mu\nu\chi}(r_k) = \frac{2}{N(N-1)V_k}\sum_{i,j:r_{ij}\in \mathrm{bin}\, k} f_m f_n D^m_{\chi\mu}(\omega_i) D^n_{\underline{\chi}\nu}(\omega_j) \tag{3.20}$$

The following data was extracted from simulation snapshots of an NVT ensemble using periodic boundary conditions (bulk system). In Figure 3.4 we show the radial distribution function $g^{lmn}_{[11]}$ (note the use of symmetry-adapted basis functions) for all non-vanishing moments with $l_{\max} \leq 4$ at a state point with strong orientational anisotropy but still above the critical temperature.

For the moments that we examine, one can see that the strongest contributions are located in the bonding region (radial distance in which two adjacent particles are bonded), however correlations also extend further out. The exact location of the second peak depends on the specific moment, but is generally before $r = 2\sigma$. This is a consequence of the particles' symmetry, leading to the formation of tetrahedral networks. It makes sense to differentiate between moments with one index $m, n = 0$, and those with $m, n \neq 0$, as the former are independent of the "central" particle's orientation. These are therefore less influenced by the tetrahedral geometry, which leads to a different structure if compared to the completely anisotropic moments. Moments with at least one index equal to zero are not as peaked around the second coordination shell (where alignment becomes important) and their range is slighlty larger.

3.4.2 Inhomogeneous case

When applying an external field the density profiles become inhomogeneous. In the following we will employ different external potentials in order to investigate the structure of the patchy fluid. We limit ourselves to potentials that are dependent only on the position $\mathbf{r}$ of the particle and not its orientation Ω. An orientation dependent external potential would lead to so-called anchoring, favoring certain orientations of the particles[2]. The first potential we are considering is the hard wall, given by

$$V_{\text{wall}}(z) = \begin{cases} \infty & \text{if } z \leq \sigma \text{ or } z \geq L \\ 0 & \text{otherwise} \end{cases} \tag{3.21}$$

which means it is independent of the x, y coordinate, and the system is assumed to extend periodically in these directions. Apart from its simplicity, the hard wall further permits the extraction of information about the pressure using the contact density, see [Rot10]. The system is realized on the computer by constraining MC moves to ensure particles remain within the box (outside the wall). In the radial

[2] DFT studies featuring an anisotropic external potential can be found in [WRS18; MCT13]

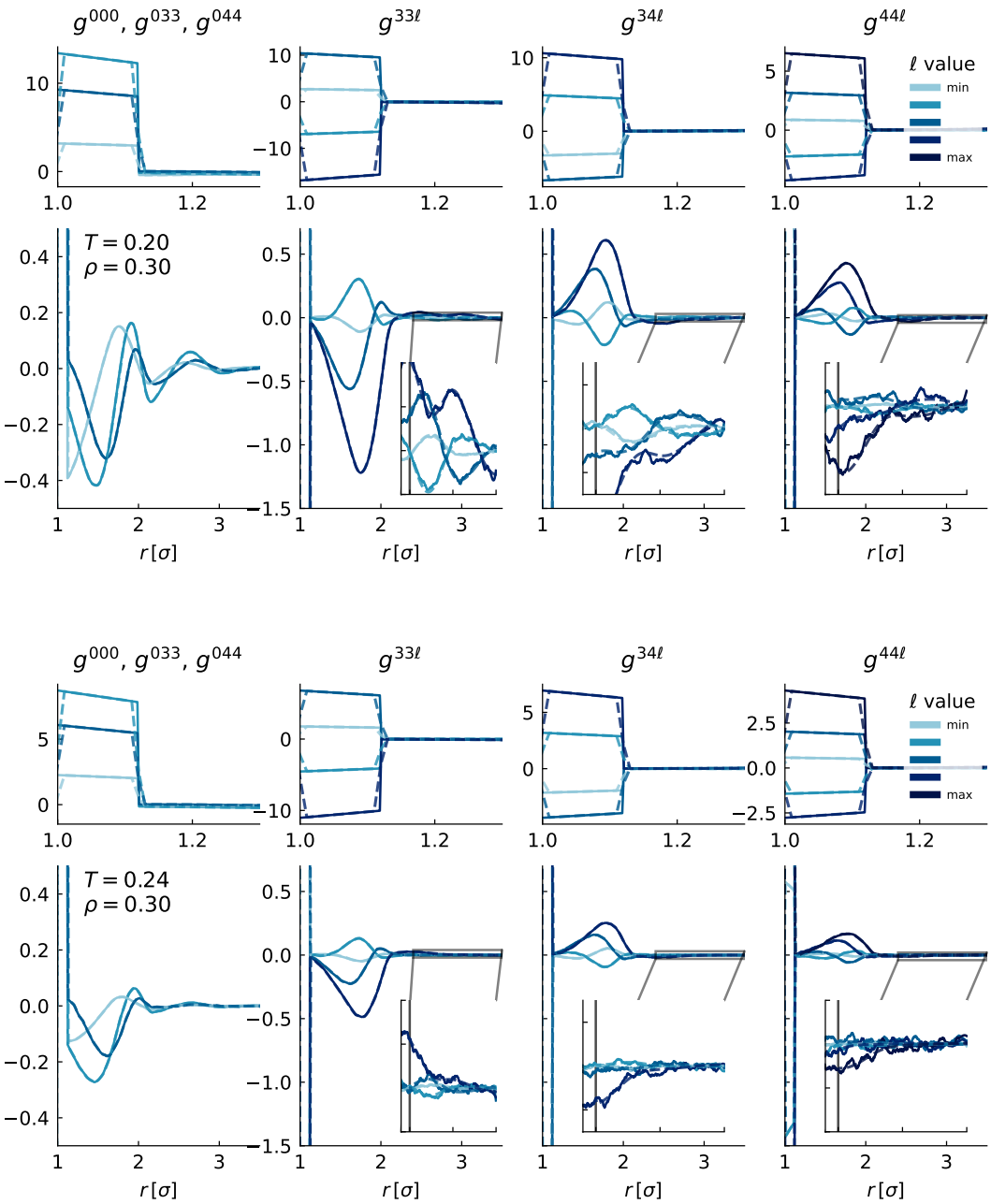
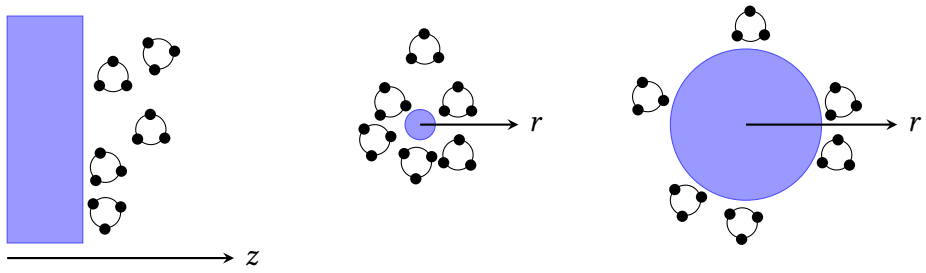


Figure 3.4: Radial distribution function $g^{mnl}_{[11]}$ for up to $l_{\max} = 4$. The upper panel shows the bonding region, while the lower panel extends to up to $r = 3.5\sigma$.



(a) (b) (c)

Figure 3.5: Diagrams representing different external potentials, shown as blue areas (exclusion volumes). Here (a) shows the hard wall, (b) and (c) show that of a hard tracer for different tracer radii. With increasing tracer radius, the system in (c) approaches that of hard wall. Intuitively, this is clear as the curvature of sphere with radius $r \to \infty$ is zero.

geometry we employ the hard tracer and the square-well (SW) potential

$$V_{\text{tracer}}(r) = \begin{cases} \infty & \text{if } r \leq \sigma \\ 0 & \text{otherwise} \end{cases} \tag{3.22}$$

$$V_{\text{sw}}(r) = \begin{cases} \infty & \text{if } r \leq \sigma \\ -\epsilon & \text{if } r > \sigma \leq \sigma + \delta \\ 0 & \text{otherwise} \end{cases} \tag{3.23}$$

Due to the radial symmetry, these profiles give us the possibility to see how the fluid structure changes when patchy particles are grouped around particles with isotropic interactions, therefore ignoring their tetrahedral geometry.

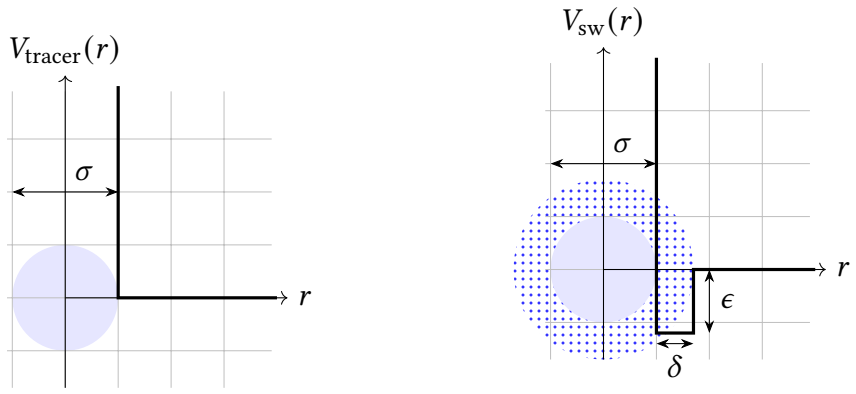
In order to examine the orientational structure of the fluid we determine the complete density distribution

$$\rho(\mathbf{r}, \Omega) = n(\mathbf{r})\alpha(\mathbf{r}, \Omega) = \sum_{lmn} n(\mathbf{r})\alpha^{l}_{mn}(\mathbf{r}, \Omega) \tag{3.24}$$

induced by the respective external potential $V(\mathbf{r})$. In the planar geometry we create histograms according to

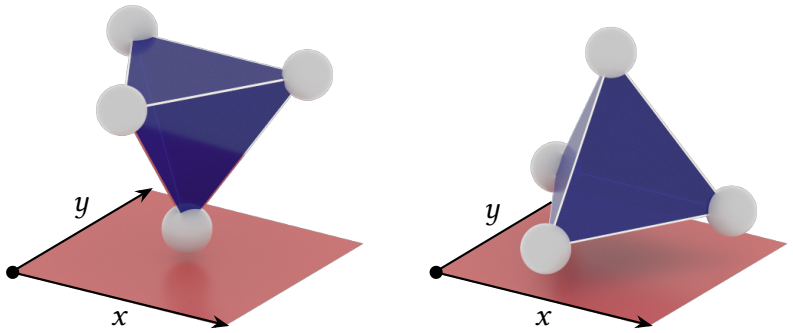
$$n(z_k) = \frac{1}{N} \sum_{i:z_i \in z_k} 1 \tag{3.25}$$

$$\alpha^{l}_{mn}(z_k) = \frac{1}{N} \sum_{i:z_i \in z_k} \overline{D^{l}_{mn}}(\Omega_i) \tag{3.26}$$



(a) (b)

Figure 3.6: Plots of the two external potentials exerted by the hard tracer (a), and the square-well particle (b) with the corresponding parameters ϵ (bonding strength) and δ (bonding range). The blue shaded areas represent the hard spherical particle and the attraction shell, respectively.



(a) (b)

Figure 3.7: The two configurations with the highest (a) and lowest (b) probability, close to the wall (shown as red square). The configuration on the l.h.s is more likely to appear, since three patches are pointing towards the fluid (away from the wall) and can engage in bonds with other particles. The wall is indicated by the red square in the xy plane.

where z_k stands for the k-th bin in the z-direction and z_i is the z-coordinate of particle i. The analysis in the radial geometry is done similarly to the one in the homogeneous case in Section 3.4.1, only that now the first loop variable is fixed on the central tracer particle, while the second iterates over all patchy particles

around it. The normalization has to be adapted to account for the fact that there is only one tracer, $N \to 1$, and only moments of $g_{\mu\nu}^{mnl}$ with $m = \mu = 0$ are non-vanishing, as the tracer is radially symmetric. Due to the symmetry rules of the $3j$-symbol, we have further the constraint $n = l$, i.e. the only moments are $g_{0\nu}^{0nn}$. Every simulation snapshot now only contains N particle pairs that can be used in the analysis (compared to $N(N-1)/2$ in the homogeneous case), so the statistics will be worse for the same amount of snapshots analyzed.

Hard wall The effect of a hard wall is dependent on the temperature. For high temperatures the hard-sphere interaction is the dominant factor, and we expect to see the usual layering structure. When lowering the temperature the bonding states become thermodynamically more favored and the orientation distribution therefore less isotropic. In addition, the effect of drying sets in, which means that the density of the fluid decreases when going from the bulk to the wall. The reason for this is that in order to maximize the number of possible bonding partners every particle needs to surround itself with as many other particles as possible. Close to the wall there are no possible bonding partners for one hemisphere of the particle and therefore this position is not energetically favorable for the particles.

Some examples of density and orientation profiles are shown in Figure 3.8. Knowledge of the individual moments still does not allow us to make direct observations on the orientational structure of the fluid. In order to gauge whether the orientational order is actually strong and which orientations are preferred, we need to look at the ODF $\alpha(z, \Omega)$.

We reconstruct the ODF α by expanding the moments α_{mn}^{l}, extracted from simulations. Since we only know them up to a certain value $l = l_{\max}$ the ODF is necessarily a truncated approximation, i.e.

$$\alpha_{\mathrm{tr}}(z, \Omega) = \sum_{lmn}^{l_{\max}} D_{mn}^{l}(\Omega)\alpha_{mn}^{l}(z) \tag{3.27}$$

Due to the orthogonality of the Wigner D-matrices, the truncated expansion ODF is always normalized, no matter at which $l_{\max}$ the expansion was truncated. Hence, we cannot use the criterion of normalization to gauge the magnitude of the neglected terms in the truncated expansion.

By plotting $K\alpha(z, \Omega)$, where $K = 8\pi^2 = \alpha_{00}^{0}$, we can directly read off the relative probabilities of certain orientational configurations, if compared to the uniform case. We show a histogram of the distribution $\alpha(z_i, \theta, \chi)$ of a KF fluid near a wall

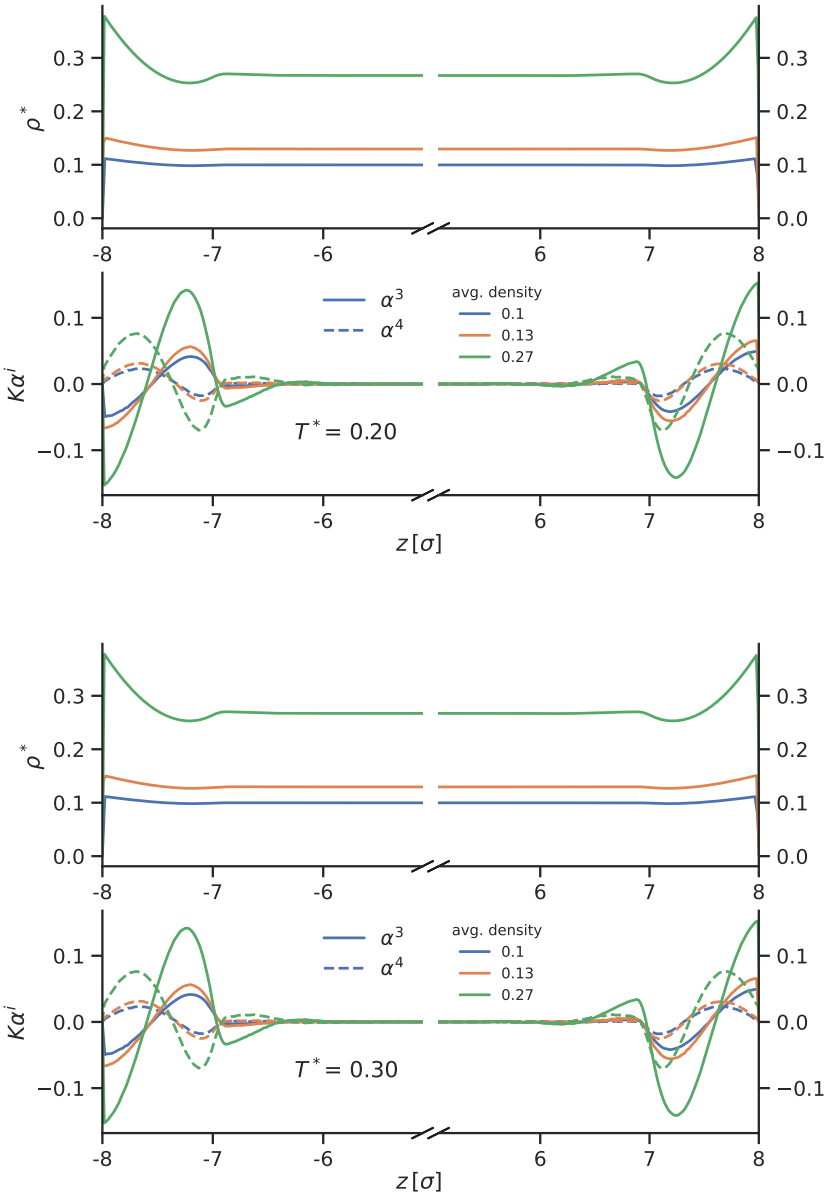


Figure 3.8: Density and orientation profiles between two hard walls for two different temperatures. The upper sub-panel shows the density profiles and the lower sub-panel the orientational profiles in the invariant basis up to $l_{\max} = 4$. Both walls are shown here to highlight the different behavior of α^3 and α^4 under change of $z \longrightarrow -z$.

in Figure 3.9. Every frame is taken at a different z_i, starting from a position just at the wall and then with increasing distance towards the bulk liquid. We can see that close to the wall the probability distribution ranges from 0.7 to about 1.3. This means that the high probability region is about twice as likely as the low probability configuration. Consequently, we can say that the anisotropy is rather strong for this state point.

Radial tracer For the radial external potentials we proceed similarly to the calculation of the radial distribution function, only that this time one "particle" is fixed by the location of the support of the external potential. We modify the code, such that in the center of the simulation box there is an extra hard-sphere potential acting on the surrounding particles. Potential MC moves are forbidden into the hard-sphere region $r \leq \sigma/2$ and the energy calculation of a trial move takes into account eventual bonding with the square well of the central particle. We show the resulting profiles in Figures 3.12 and 3.13, for different values of ρ^*_{avg}, T^* and square-well depth ϵ. In addition, we show the density profiles resulting from the Stopper–Wu (StWu) functional[3].

Starting with the case $\epsilon = 0$, we can see that the position of the second coordination shell is underestimated. Increasing the bonding energy of the square-well tracer, one observes that the contact densities are severely underestimated as well. The functional predicts fewer neighbors than observed in the simulation.

If we compare to the RDFs shown in Figure 3.4, we see that there, the second peak occurs earlier than in Figures 3.12 and 3.13, a consequence of the formation of tetrahedral networks. The isotropic tracer particle hinders the formation of such networks in the first coordination shell and as a consequence the second peak is located at about $r \approx 2\sigma$, in line with the behavior of a hard-sphere liquid.

From simulation snapshots we can observe that the patchy particles are oriented around the tracer similar to how they were in front of the hard wall, i.e. with the patch-poor region pointed towards the tracer. We can also make some sense of the orientational density distributions shown in Figures 3.12 and 3.13. While the density distributions $\rho^{0nn}_{0\nu}(\Omega, \mathbf{r}) = \rho g^{0nn}_{0\nu}(\Omega, \mathbf{r})$ do formally depend on both the patchy particle's orientation and the connection vector $\mathbf{r}$, they must be independent of the direction $\mathbf{r}$ due to the spherical symmetry of the tracer potential. Indeed, by

[3]The functional will be introduced in more detail later. It will serve as our reference functional, which makes the comparison relevant. Furthermore, we show in Figure 3.10 a comparison of the density profiles generated using the Stopper–Wu functional and simulation data between two hard walls. No orientations are shown here, as the functional assumes isotropy of the liquid.

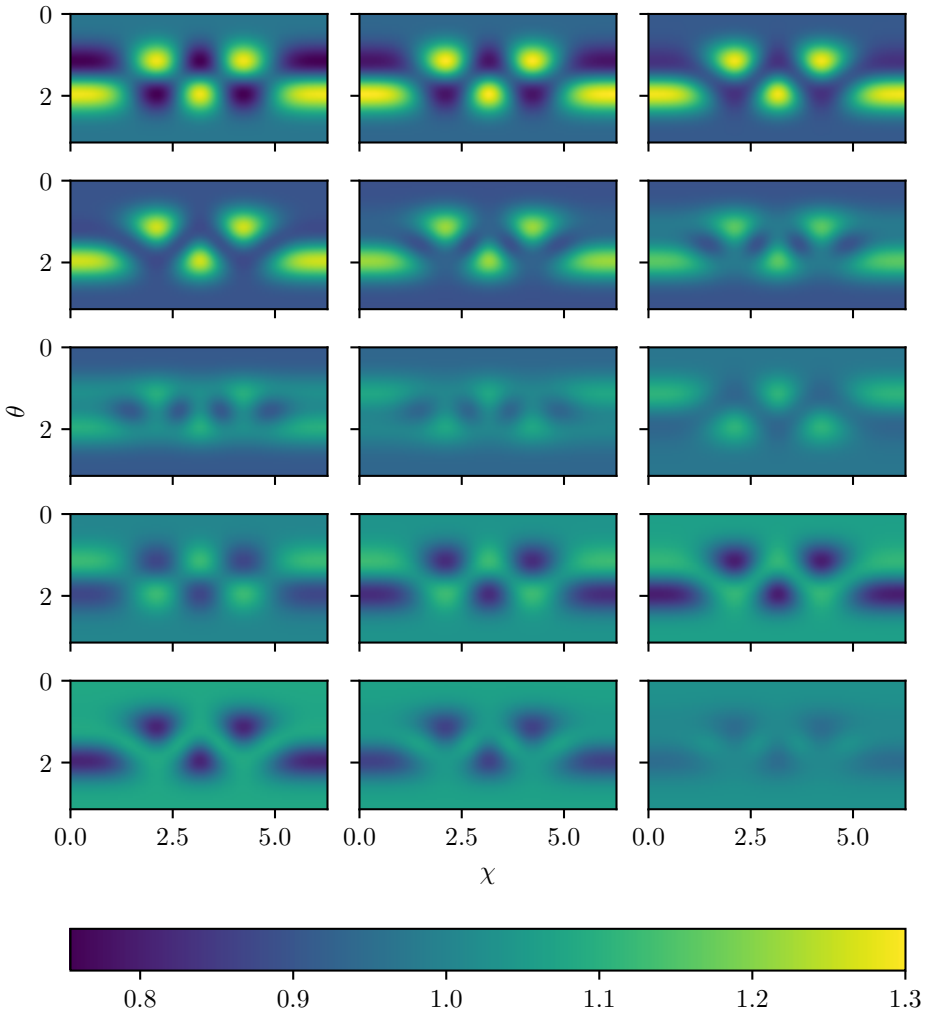


Figure 3.9: Orientational distribution function $\alpha(z_i, \theta, \chi)$ for different distances z_i from the wall (starting at distance zero and then increasing by 0.072σ in every step) at the state point $T^* = 0.2$, $\rho^* = 0.3$. Reading direction in increasing distance from the wall is as follows:  , i.e. top to bottom, left to right. Close to the wall, we see a strong ordering effect, indicated by the fourfold (due to tetrahedral symmetry) appearance of high-probability and low-probability configurations. With growing distance to the wall the low-probability configurations start to vanish, leaving only certain favored high-probability configurations. Further out, the situation reverses (low ↔ high), before finally becoming relatively uniform.

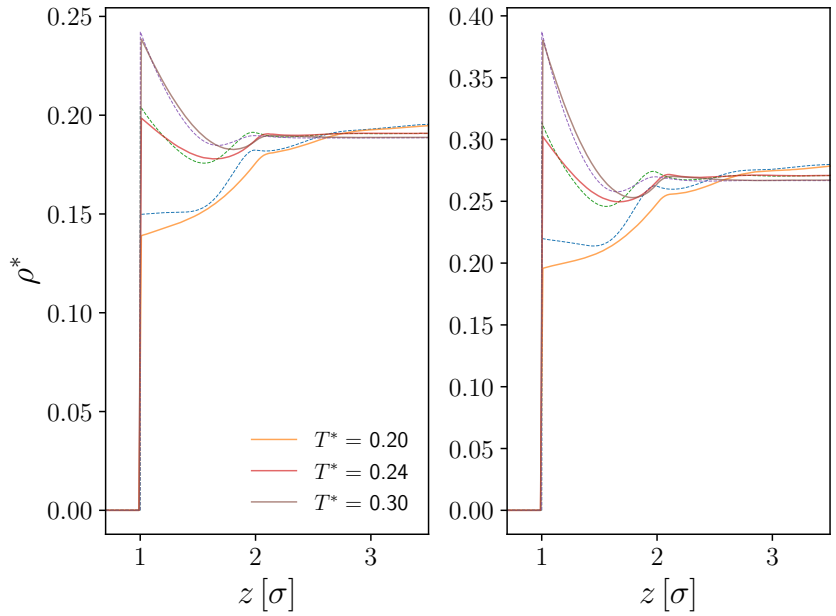
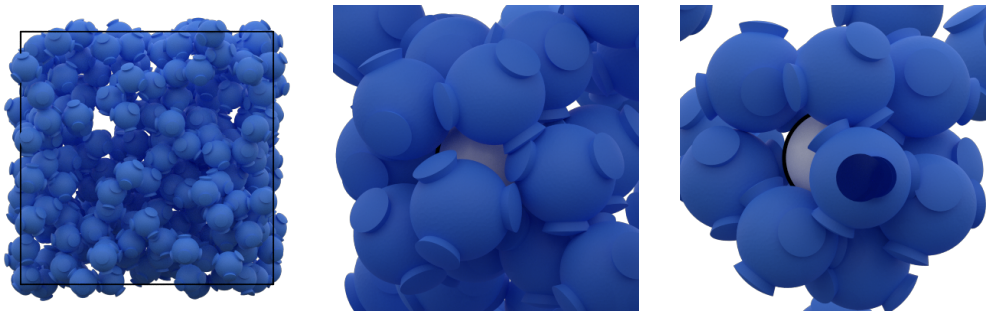


Figure 3.10: Comparison of density profiles generated using the Stopper–Wu functional (dashed line) and the MC data (full line) for the KF potential. For low temperatures it is evident that the shift of the position of the first peak is not accounted for. Further, drying at the wall is underestimated, which means that the EOS is also not accurately captured (due to equivalence between the contact value and the pressure).



(a) (b) (c)

Figure 3.11: Simulation snapshots of the patchy particles under the influence of different external potentials. In (a) we show the results for a hard wall potential in z-direction, looking into the xy-plane from the wall. In (b) and (c) the tracer is shown as a white particle with black outline. For both situations (hard wall and hard tracer), the particles align such that they point most of the patches into the direction of neighboring patchy particles.

performing the χ-transform on the moments (i.e. changing into the intermolecular frame), this yields moments that only depend on the orientation Ω and the radial distance r. By using Equation (2.166) we get

$$g^{0n}_{0\nu\chi} = \begin{pmatrix} 0 & n & n \\ \chi & \underline{\chi} & 0 \end{pmatrix} g^{0nn}_{0\nu} \tag{3.28}$$

Due to the symmetries of the $3j$-symbol, only the term with $\chi = 0$ is non-zero. Expanding the IM distribution function, according to Equation (2.168) we get

$$g(r,\omega) = \sum_{n\nu} g^{0n}_{0\nu 0}\, \overline{D^{n}_{0\nu}}(\omega) \tag{3.29}$$

We can now compare g^{03}_{020} and g^{04}_{000} to the quantities α^3 and α^4, computed at the hard wall, cf. Figure 3.8. We need to evaluate two $3j$-symbols, i.e.

$$\begin{pmatrix} 0 & 3 & 3 \\ 0 & 0 & 0 \end{pmatrix} = -\sqrt{1/7}, \qquad \begin{pmatrix} 0 & 4 & 4 \\ 0 & 0 & 0 \end{pmatrix} = 1/3 \tag{3.30}$$

and note that $\Delta^4 \propto -D^4_{00} \propto D^4_{04}$. This means that $g(r,\omega)$ is dominated by $D^3_{02}(\omega)$, while $\alpha(z,\Omega)$ is dominated by $-D^3_{02}(\Omega)$, close to the wall. The $l = 4$ moments similarly have this inverse behavior. In other words, the orientation distribution around the tracer is proportional to the distribution in front of the hard wall at the

second coordination shell, because it is here where the signs of the expansion of α^3, α^4 change. Thus, it is exactly opposite to the distribution in front of the hard wall (first coordination shell), seen in the first panel in Figure 3.9. The reason for this is, that in order to maximize the number of possible bonding partners in front of the wall, the best orientation is with one of the patches towards the wall. This way the remaining three patches are all aligned in a plane parallel to the wall (i.e. the xy-plane). Around the tracer the situation is slightly different, here the particles surround the tracer radially and can form more bonds if the side opposite to one of the patches, points towards the tracer. This can be observed nicely in panel (b) of Figure 3.11.

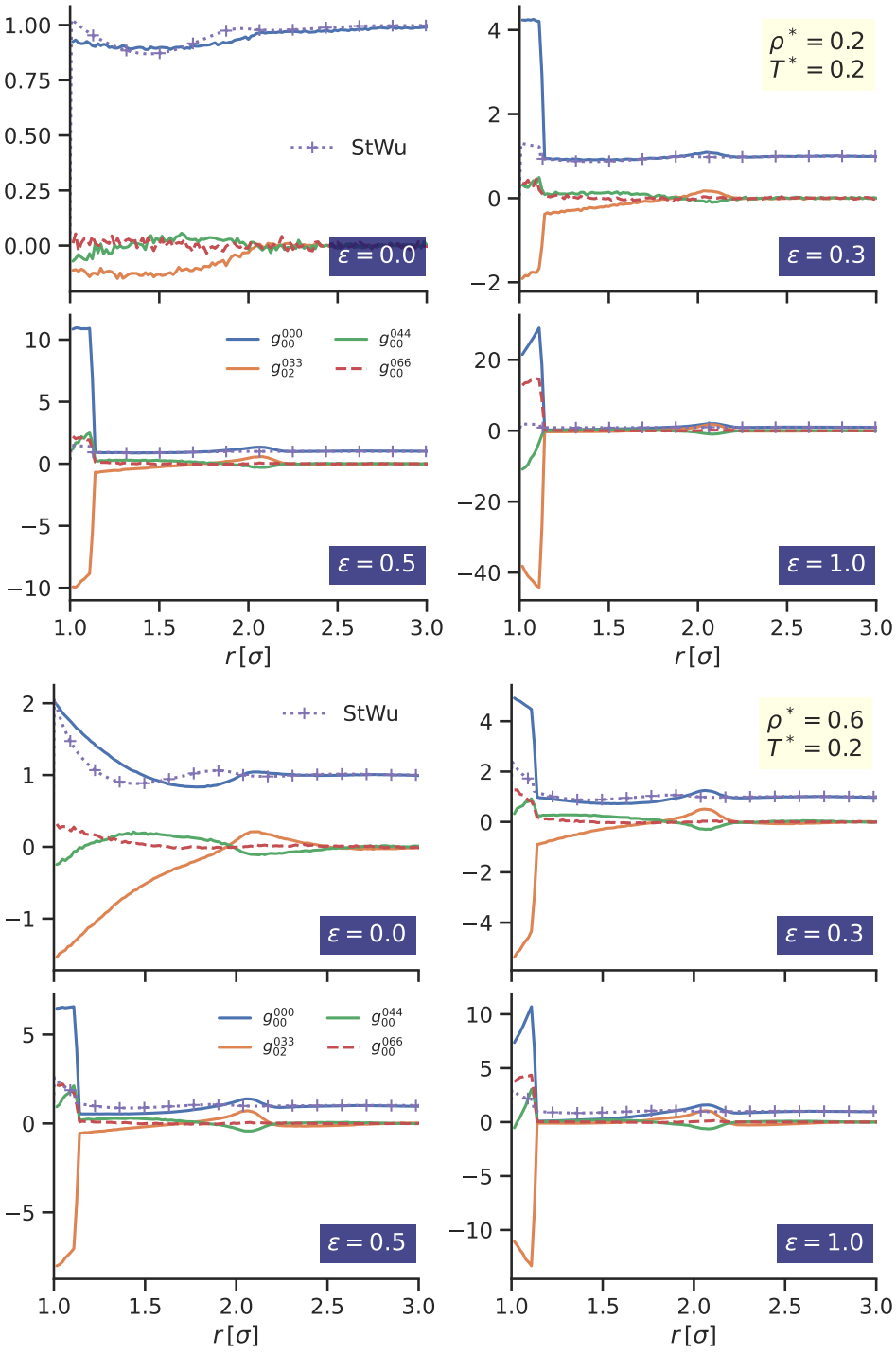


Figure 3.12: Density profiles of patchy particles around radial tracers with different attraction strengths ϵ at low temperature T.

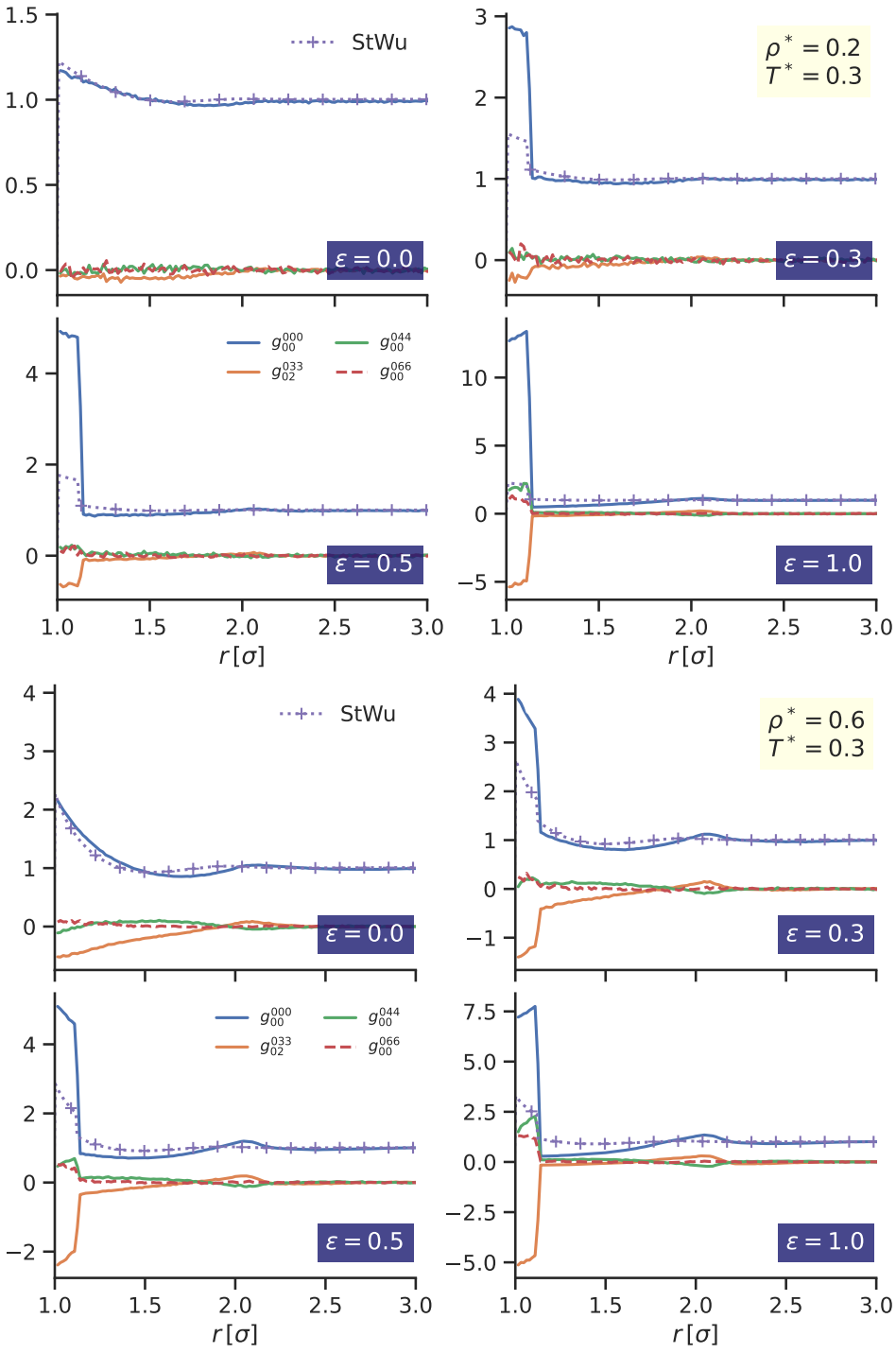


Figure 3.13: Density profiles of patchy particles around radial tracers with different attraction strengths ϵ at high temperature T.

4 Towards a density functional theory treatment

> Tyger Tyger, burning bright
> In the forests of the night
> What immortal hand or eye
> Could frame thy fearful symmetry?
>
> — William Blake, *The Tyger*

In the following we present the necessary equations for density functional treatment of an anisotropic system. The main difference w.r.t. the isotropic case is the additional Euler–Lagrange (EL) equation that leads to the equilibrium orientation distribution function. Further, the EL equation for the density profile gains a term that is related to entropy of the ODF α.

4.1 The DFT formalism for anisotropic densities

The main obstacle of density functional theory is finding an accurate analytical expression for the grand potential functional $\Xi[\rho]$, which is a functional of the density profile $\rho(\mathbf{r}, \Omega)$. The equilibrium state is then found by minimization of the grand potential functional

$$\frac{\delta\beta\Xi[\rho]}{\delta\rho(\mathbf{r}, \Omega)} = 0\,, \quad \text{for } \rho = \rho_{\text{eq}}(\mathbf{r}, \Omega) \tag{4.1}$$

The grand potential can be written as

$$\beta\Xi = \beta\mathcal{F}[\rho(\mathbf{r}, \Omega)] + \int \mathrm{d}\mathbf{r}\, \mathrm{d}\Omega \rho(\mathbf{r}, \Omega) \left[\beta V^{\text{ext}}(\mathbf{r}, \Omega) - \beta\mu\right] \tag{4.2}$$

with the Helmholtz free energy $\mathcal{F}$, the external potential V^{ext} and the chemical potential μ. One can further split the free energy into an ideal, non-interacting term $\mathcal{F}_{\text{id}}$ and the excess term $\mathcal{F}_{\text{ex}}$ that results from particle interactions

$$\mathcal{F}[\rho(\mathbf{r},\Omega)] = \mathcal{F}_{\text{id}}[\rho(\mathbf{r},\Omega)] + \mathcal{F}_{\text{ex}}[\rho(\mathbf{r},\Omega)] \ . \tag{4.3}$$

The ideal gas term is given by

$$\beta\mathcal{F}_{\text{id}} = \int \mathrm{d}\mathbf{r}\,\mathrm{d}\Omega\rho(\mathbf{r},\Omega)\left[\ln(\rho\Lambda^3 K) - 1\right] \tag{4.4}$$

and the second, excess term, is in general unknown. Here Λ is the thermal de-Broglie length and K is either $8\pi^2$ for a general molecule or 4π, in the case of an axially symmetric one. In order to separate the free energy contributions that are due to the orientation of the particles from the orientation-independent effects we split $\rho(\mathbf{r},\Omega)$ into an orientationally averaged component $n(\mathbf{r})$ (depending only on $\mathbf{r}$) and an orientation distribution function $\alpha(\mathbf{r},\Omega)$

$$\rho(\mathbf{r},\Omega) = n(\mathbf{r})\alpha(\mathbf{r},\Omega) \tag{4.5}$$

with

$$\int \mathrm{d}\Omega\alpha(\mathbf{r},\Omega) = 1 \tag{4.6}$$

The ideal part of the Helmholtz free energy in the anisotropic case becomes

$$\begin{aligned}
\beta\mathcal{F}_{\text{id}} &= \int \mathrm{d}\mathbf{r}\,\mathrm{d}\Omega n(\mathbf{r})\alpha(\mathbf{r},\Omega)\left[\ln(n(\mathbf{r})\alpha(\mathbf{r},\Omega)\Lambda^3 K) - 1\right] \\
&= \int \mathrm{d}\mathbf{r}\,\mathrm{d}\Omega n(\mathbf{r})\alpha(\mathbf{r},\Omega)\left[\ln(n(\mathbf{r})\Lambda^3) - 1 + \ln(K\alpha(\mathbf{r},\Omega))\right] \\
&= \int \mathrm{d}\mathbf{r} n(\mathbf{r})\left[\log(\Lambda^3 n(\mathbf{r})) - 1\right] + \int \mathrm{d}\mathbf{r}\,\mathrm{d}\Omega n(\mathbf{r})\alpha(\mathbf{r},\Omega)\log(K\alpha(\mathbf{r},\Omega))
\end{aligned} \tag{4.7}$$

The excess free energy can then in general be written as

$$\mathcal{F}_{\text{ex}} = \mathcal{F}_{\text{ex,iso}}[n(\mathbf{r})] + \mathcal{F}_{\text{ex,ori}}[n(\mathbf{r}),\alpha(\mathbf{r},\Omega)] \tag{4.8}$$

where we expect that some parts of the free energy $\mathcal{F}_{\text{ex,iso}}$ are entirely determined by the fluid's density alone, such as the hard-core interaction, while others will probably also need information about the orientation of the particles. Anisotropic interactions, of which the angular dependence was integrated out also fit into the first category.

Next, we need to derive the Euler–Lagrange equations w.r.t. the density and the orientation. The normalization of α leads to the inclusion of a Lagrange multiplier λ. We therefore have

$$\frac{\delta\beta\Xi}{\delta n(\mathbf{r})} = 0 \tag{4.9}$$

$$\frac{\delta\beta\Xi}{\delta\alpha(\mathbf{r},\Omega)} = \lambda \tag{4.10}$$

For convenience we also write out Ξ again

$$\begin{aligned}\beta\Xi = &\int \mathrm{d}\mathbf{r} n(\mathbf{r}) \left[\log(\Lambda^3 n(\mathbf{r})) - 1\right] + \int \mathrm{d}\mathbf{r}\,\mathrm{d}\Omega n(\mathbf{r})\alpha(\mathbf{r},\Omega)\log(K\alpha(\mathbf{r},\Omega)) \\ &+ \beta\mathcal{F}_{\mathrm{ex}} + \int \mathrm{d}\mathbf{r}\,\mathrm{d}\Omega\rho(\mathbf{r})\alpha(\mathbf{r},\Omega)\beta V^{\mathrm{ext}}(\mathbf{r},\Omega) - \int \mathrm{d}\mathbf{r} n(\mathbf{r})\beta\mu\end{aligned} \tag{4.11}$$

The density part The first derivative becomes

$$\begin{aligned}0 &= \ln(\Lambda^3 n(\mathbf{r})) + \int \mathrm{d}\Omega\alpha(\mathbf{r},\Omega)\ln(K\alpha) + \frac{\delta\beta\mathcal{F}_{\mathrm{ex}}}{\delta n(\mathbf{r})} + \int \mathrm{d}\Omega\alpha(\mathbf{r},\Omega)\left[\beta V^{\mathrm{ext}} - \beta\mu\right] \\ &= \ln(\Lambda^3 n(\mathbf{r})) + \langle\ln(K\alpha)\rangle_\alpha(\mathbf{r}) + \frac{\delta\beta\mathcal{F}_{\mathrm{ex}}}{\delta n(\mathbf{r})} + \beta V^{\mathrm{ext}}(\mathbf{r}) - \beta\mu\end{aligned} \tag{4.12}$$

where we wrote $\int \mathrm{d}\Omega\alpha(\mathbf{r},\Omega)X = \langle X\rangle_\alpha$ and used the fact that α is normalized and that $V^{\mathrm{ext}}(\mathbf{r})$ should not depend on the orientation. Solving this equation for $n(\mathbf{r})$ we have

$$n(\mathbf{r}) = \exp\left[-\langle\ln K\alpha\rangle_\alpha(\mathbf{r}) - \frac{\delta\beta\mathcal{F}_{\mathrm{ex}}}{\delta n(\mathbf{r})} - \beta V^{\mathrm{ext}}(\mathbf{r}) + \beta\mu^*\right] \tag{4.13}$$

with

$$\beta\mu^* = \beta\mu - \ln\Lambda^3 \tag{4.14}$$

The orientational part The second term is given by

$$\lambda = n(\mathbf{r})\ln(K\alpha(\mathbf{r},\Omega)) + n(\mathbf{r}) + \frac{\delta\beta\mathcal{F}_{\mathrm{ex,ori}}}{\delta\alpha(\mathbf{r},\Omega)} + n(\mathbf{r})\beta V^{\mathrm{ext}}(\mathbf{r},\Omega) \tag{4.15}$$

which, solved for α becomes (assuming $\rho > 0$)

$$\begin{aligned}\alpha(\mathbf{r},\Omega) &= K^{-1}\exp\left[\frac{\lambda - n(\mathbf{r})}{n(\mathbf{r})} - \frac{1}{n(\mathbf{r})}\frac{\delta\beta\mathcal{F}_{\mathrm{ex,ori}}}{\delta\alpha(\mathbf{r},\Omega)} - \beta V^{\mathrm{ext}}(\mathbf{r},\Omega)\right] \\ &= \frac{1}{K}\exp\left[\frac{\lambda - n(\mathbf{r})}{n(\mathbf{r})}\right]\psi(\mathbf{r},\Omega)\end{aligned} \tag{4.16}$$

where we defined the unnormalized distribution function as

$$\psi(\mathbf{r}, \Omega) = \exp\left[-\frac{1}{n(\mathbf{r})}\frac{\delta\beta\mathcal{F}_{\text{ex,ori}}}{\delta\alpha(\mathbf{r}, \Omega)} - \beta V^{\text{ext}}(\mathbf{r}, \Omega)\right] \tag{4.17}$$

The normalization conditions implies that

$$\frac{\lambda - n(\mathbf{r})}{n(\mathbf{r})} = \ln\left[\frac{K}{\int \mathrm{d}\Omega\psi(\mathbf{r}, \Omega)}\right] \tag{4.18}$$

Inserting this into Equation (4.16) results in the following set of equations

$$n(\mathbf{r}) = \exp\left(-\langle \ln K\alpha\rangle_\alpha(\mathbf{r}) + \beta\mu - \beta V^{\text{ext}}(\mathbf{r}) - \frac{\delta\beta\mathcal{F}_{\text{ex}}[n, \alpha]}{\delta n(\mathbf{r})}\right) \tag{4.19}$$

$$\begin{aligned}\alpha(\mathbf{r}, \Omega) &= Z^{-1}(\mathbf{r})\psi(\mathbf{r}, \Omega) \\ &= Z^{-1}(\mathbf{r}) \exp\left(-\frac{1}{n(\mathbf{r})}\frac{\delta\beta\mathcal{F}_{\text{ex,ori}}[n, \alpha]}{\delta\alpha(\mathbf{r}, \Omega)} - \beta V^{\text{ext}}(\mathbf{r}, \Omega)\right) \qquad (4.20)\end{aligned}$$

$$Z(\mathbf{r}) = \int \mathrm{d}\Omega\psi(\mathbf{r}, \Omega) \tag{4.21}$$

4.2 Previous treatments of anisotropic fluids

Classical density functional theory has been applied to the study of associating fluids before. One of the first papers that combined the DFT formalism with anisotropic interactions was [CTN83], where the authors consider the structure of a liquid–vapor interphase of a diatomic molecule (CO), interacting through an (isotropic) Lennard-Jones potential and an additional anisotropic potential resulting from dipole–quadrupole and quadrupole–quadrupole interactions. Later the authors of [TG91] investigate the structure of spherical particles interacting through a dipole–dipole interaction.

The first time a concrete model for water was considered is in [Yan+91; Yan+92], using an extended mean-field approach. Again the liquid–vapor interface is considered.

Multiple authors cover the treatment of a classical Heisenberg fluid [Tav+95; SGK14; CGS16; WRS18]. Here, the particles interact through a spin–spin coupling term, which has a different orientational dependence than a dipole–dipole interaction.

A study of bivalent patchy particles, taking into account also the orientations can be found in [TS19]. Because the particles have only two patches, they are symmetric around the z-axis, simplifying the treatment somewhat, as all orientational quantities can be expanded in spherical harmonics instead of Wigner D-matrices. Similar to our approach, Wertheim's TPT1 serves as the reference functional for the isotropic bonding interaction. The hard sphere part was modelled with a weighted density approximation functional, which was observed to perform more accurately than FMT in the case of patchy particles with three patches. The anisotropic bonding addition takes the form of a mean-field approximation. The approach lead to density and orientation profiles that agree reasonably well with simulations at high temperatures but fail for lower temperatures.

4.3 Wertheims thermodynamic perturbation theory

4.3.1 Original formulation

It is clear that the strong directionality of associating liquids needs to be taken into account in order to make sufficiently precise predictions. In the *thermodynamic perturbation theory* (TPT), developed by M. S. Wertheim [Wer84a; Wer84b; Wer86a; Wer86b] around forty years ago, this is done on the level of the cluster expansion of the free energy. The theory assumes a hard sphere reference liquid, with colloids decorated by short-range attraction sites. The exact geometrical details are not important, in fact the theory developed to first order (TPT1) depends on the bonding volume alone and no further information is needed. In order to be able to reduce the number of relevant terms in the expansion, a number of approximations have to be made. These are: the single bond per patch constraint, ignoring the formation of ring structures and the assumption that the correlation of the unbonded colloids is equal to the one of the reference liquid. The complete calculation is quite involved, but a detailed derivation can be found [ZG16]. Additional information regarding the approximation that are made in the derivation can be found [Mar14; MC16]. An important insight of Wertheim was that the treatment could be simplified by treating particles in different bonding states (monomers, dimers or in general s-mers, see Figure 4.1) as different species, in the so-called multi-density formalism. In this case, the total density can be separated into the density of non-bonded monomers ρ_0 and the bonded dimers $\rho - \rho_0$. Then the resulting expression for the

(a) monomer (b) dimer

Figure 4.1: Monomers and dimers from Wertheims original formulation. The circles represent colloids with a single bonding site.

free energy reads

$$\frac{F_{\text{bond}}}{Nk_BT} = \ln X - \frac{1}{2}X + \frac{1}{2}, \tag{4.22}$$

where $X = \rho_0/\rho$ is the monomer fraction. It is determined by the solution of the equation

$$X^2\rho_b\Delta + X - 1 = 0 \tag{4.23}$$

where Δ is given by

$$\Delta = \int g_{\text{hs}}(\mathbf{r}) \left\langle e^{-\beta\phi_{\text{bond}}} \right\rangle \mathrm{d}\mathbf{r} \tag{4.24}$$

where $\langle\cdots\rangle$ denotes the angular average and $g_{\text{hs}}(\mathbf{r})$ is the RDF of the reference hard-sphere system. Approximating it by the contact value $g_{\text{hs}}(\sigma^+)$ we can write

$$\Delta \approx v_b g_{\text{hs}}(\sigma^+) \left(e^{\beta\epsilon} - 1\right) \tag{4.25}$$

The previous result only applies to particles with one patch each. It was later generalized by Wertheim to particles with multiple patches [Wer86a] sites $A, B, \ldots \in \Gamma$, and reads in the notation of [JCG88]

$$\frac{F_{\text{bond}}}{Nk_BT} = \sum_{A\in\Gamma} \left(\ln X_A - \frac{1}{2}X_A\right) + M\frac{1}{2} \tag{4.26}$$

Instead of the unbonded monomer fraction X, one has now X_A, the fraction of particles *not* bonded at site A, determined by

$$X_A = \frac{1}{1 + \rho \sum_{B\in\Gamma} X_B\Delta_{AB}} \tag{4.27}$$

with Δ_{AB} given by

$$\Delta_{AB} = \int g_{\text{hs}}(\mathbf{r}) \left\langle e^{-\beta\phi_{\text{AB}}} \right\rangle \mathrm{d}\mathbf{r} \tag{4.28}$$

In our case, we treat particles with equivalent bonding sites, i.e. they all share the same interaction potential, and they all can interact with each other. In this case the quantities reduce to $\phi_{\mathrm{AB}} = \phi_{\mathrm{patch}}$, and $\Delta_{AB} = \Delta$. Then the free energy becomes

$$\frac{F_{\mathrm{bond}}}{Nk_BT} = M\left(\ln X - \frac{1}{2}X + \frac{1}{2}\right) \tag{4.29}$$

and the equation that determines X becomes

$$X = \frac{1}{1 + MX\rho\Delta} \tag{4.30}$$

4.3.2 Generalizations to the inhomogeneous case

Starting from this bulk theory of associating liquids, there were multiple attempts to generalize this to the inhomogeneous case. Early applications by Segura et al. [SCS97], utilized local density and weighted density approximations, with the WDA leading to better results. A straightforward functionalization of Equation (4.26) would result in the free energy density

$$\Phi_{\mathrm{bond}} = M\rho(\mathbf{r})\left[\ln X(\mathbf{r}) - \frac{X(\mathbf{r})}{2} + \frac{1}{2}\right] \tag{4.31}$$

where every occurrence of ρ in the determination of X and Δ is replaced by $\rho(\mathbf{r})$, and in the WDA approach replacing ρ by some averaged $\overline{\rho}(\mathbf{r})$. As a result, the quantities X, Δ also become position dependent. Later approaches by Yu and Wu [YW02a] used the framework of FMT (i.e. the weighted densities) to generalize the homogeneous functional to

$$\Phi_{\mathrm{bond}}^{\mathrm{YuWu}} = Mn_0(\mathbf{r})\xi(\mathbf{r})\left[\ln X(\mathbf{r}) - \frac{X(\mathbf{r})}{2} + \frac{1}{2}\right] \tag{4.32}$$

$$X(\mathbf{r}) = \frac{1}{1 + Mn_0(\mathbf{r})\xi(\mathbf{r})X(\mathbf{r})\Delta(\mathbf{r})} \tag{4.33}$$

introducing the new quantity $\xi = 1 - \mathbf{n}_2(\mathbf{r}) \cdot \mathbf{n}_2(\mathbf{r})/n_2^2(\mathbf{r})$. Note that in the homogeneous limit, $\xi \to 1$ and due to $\int \mathrm{d}\mathbf{r} w_0(\mathbf{r}) = 1$, $n_0(\mathbf{r}) \to \rho$ leaving the bulk theory unchanged. Additions by Stopper et al. [Sto+18] further improved the version of Yu and Wu by introducing a phenomenological parameter q, to arrive at the free energy $\Phi_{\mathrm{bond}}^{\mathrm{StWu}}$, termed by the authors modified FMT (mFMT). For reference, we

write out the complete expression, which reads

$$\Phi_{\text{bond}}^{\text{StWu}} = M n_0(\mathbf{r}) \xi^q(\mathbf{r}) \left[\ln X(\mathbf{r}) - \frac{X(\mathbf{r})}{2} + \frac{1}{2} \right] \tag{4.34}$$

$$\frac{1 - X(\mathbf{r})}{X^2(\mathbf{r})} = n_0(\mathbf{r}) \xi^q(\mathbf{r}) M \Delta(\{n_\alpha(\mathbf{r})\}) \tag{4.35}$$

$$\Delta(\{n_\alpha(\mathbf{r})\}) = v_b \left(e^{\beta\epsilon} - 1 \right) g_{\text{hs}}(\{n_\alpha(\mathbf{r})\}) \tag{4.36}$$

$$g_{\text{hs}}(\{n_\alpha(\mathbf{r})\}) = \frac{1}{1 - n_3} + \frac{\sigma n_3 \xi^q}{4(1 - n_3)^2} + \frac{\sigma^2 n_2^2 \xi^q}{72(1 - n_3)^3} \tag{4.37}$$

It was found by Stopper et al. that a value of $q = 3$ leads to most accurate results.

In the analysis [Sto19] two different external potentials were considered. First that of a hard radial tracer. While the StWu functional (mFMT) performed better than the YuWu (FMT) or WDA functional, especially with respect to contact densities and structural intensity (amplitude of density modulations), it could not predict correctly the location of the coordination shells. The radially symmetric tracer particle leads to the breaking of tetrahedral networks in the first coordination shell, which as a result is located at $r \approx 2\sigma$. Instead, the StWu functional predicts a location of $r \approx 1.7\sigma$.

The second case was that of a radial square-well tracer. Here the external potential was used to produce accurate "radial distribution functions". Strictly speaking, the density profile of patchy particles around a SW particle does not constitute the physical RDF, but by tuning the well depth ϵ_{sw} of the square well potential one can produce "effective RDFs" that come rather close to simulation results. For this to work, the parameter ϵ_{sw} is tuned such that the contact values inside the bonding region agree. It turns out that for this to be the case the parameter range is $\epsilon_{\text{sw}}/\epsilon \approx 0.25$ to 0.35.

Even considering some of these shortcomings, the StWu functional is a valuable reference since it gives a rather good approximation to the bulk equation of state (especially for medium high temperatures, see also the contact densities in Figure 3.10) and an almost quantitative agreement for the density profiles in some inhomogeneous situations. It is therefore sensible to add another perturbative contribution that takes care of predicting the orientational dependence accurately. Similar to before it comes natural to make a mean-field ansatz.

4.4 Correctional mean-field ansatz

Our approach is based on the decomposition

$$\mathcal{F}_{\mathrm{ex}}[\rho,\alpha] = \mathcal{F}_{\mathrm{id}}[\rho,\alpha] + \mathcal{F}_{\mathrm{hs}}[\rho] + \mathcal{F}_{\mathrm{bond}}[\rho] + \mathcal{F}_{\mathrm{mf}}[\rho,\alpha] \tag{4.38}$$

Strictly speaking this would not correspond to a "true" mean-field ansatz because the bonding term is already partly accounted for by the Stopper–Wu functional $\mathcal{F}_{\mathrm{bond}}$. The reasoning behind this ansatz is that the isotropic functionals already deliver qualitatively accurate results, hence we want to keep them as a reference and only learn the "corrections" in a mean-field fashion.

In order to account for the additional angle-dependent interaction between the particles we make a simple quadratic ansatz in the orientational moments, which interact through a general potential V. One prescription would be to set V equal to the interparticle interaction potential. This leads to the so-called random phase approximation (RPA), which we will cover in more detail later. In general however one can choose V freely. The ansatz looks then like

$$\begin{aligned}\mathcal{F}_{\mathrm{mf}} = \frac{1}{2}\int \mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2\,\mathrm{d}\Omega_1\,\mathrm{d}\Omega_2 \rho(\mathbf{r}_1,\Omega_1)\alpha(\mathbf{r}_1,\Omega_1)\\ \times \rho(\mathbf{r}_2,\Omega_2)\alpha(\mathbf{r}_2,\Omega_2)V(\mathbf{r}_1-\mathbf{r}_2,\Omega_1,\Omega_2)\end{aligned} \tag{4.39}$$

Note that defined in this way, V is a dimensionful quantity (energy). One could alternatively define it $\mathcal{F}$ such that the r.h.s. includes a factor k_BT making it dimensionless, however we will not make use of this. Next, we expand all quantities with angular dependence into the symmetrized basis functions,

$$\alpha(\mathbf{r},\Omega) = \sum_{b\beta' j} \alpha^{b}_{\beta'[j]}(\mathbf{r})\Delta^{b}_{\beta'[j]}(\Omega) \tag{4.40}$$

and

$$V^{\mathrm{MF}}(\mathbf{r},\Omega_1,\Omega_2) = \sum_{mnl[ij]} M^{mnl}_{[ij]}(r)\Phi^{mnl}_{[ij]}(\Omega_1,\Omega_2,\hat{\mathbf{r}}) \tag{4.41}$$

Expanding the rotational invariant basis function and taking advantage of their orthogonality we arrive at

$$\begin{aligned}\mathcal{F}_{\mathrm{ex}} = \frac{1}{2}\int \mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 n(\mathbf{r}_1)n(\mathbf{r}_2)\sum_{AB}\alpha^{a}_{\alpha'[i]}(r_1)\alpha^{b}_{\beta'[j]}(r_2)\\ \times\left[\sum_{l\lambda'} M^{abl}_{[ij]}(r_{12})\begin{pmatrix} a & b & l\\ \alpha' & \beta' & \lambda'\end{pmatrix}\overline{D^{l}_{\lambda' 0}}(\hat{\mathbf{r}}_{12})\right]\frac{64\pi^4}{c_a c_b}\end{aligned} \tag{4.42}$$

This expression is the general three-dimensional results of the mean-field ansatz written in the spherical expansion. We can simplify it further by assuming certain external symmetries. In our case, we are interested in planar symmetry, i.e. the external potential depends only on the Cartesian variable z. One consequence is that the particles orientational configuration becomes independent of the last (Euler) rotation around the z-axis, which in the spherical expansion means that the only surviving expansion coefficient corresponding to this last rotation is that of zero $\alpha^b_{\beta'[j]} \rightarrow \alpha^b_{0[j]}$

$$\mathcal{F}_{\text{ex}} = \frac{A}{2} \int \mathrm{d}z\, \mathrm{d}z' n(z) n(z') \sum_{i'j'} \alpha^{i'}(z) \alpha^{j'}(z') M^{i'j'}(z - z') \tag{4.43}$$

with

$$M^{i'j'}(z) = \sum_l \begin{pmatrix} a & b & l \\ 0 & 0 & 0 \end{pmatrix} \int \mathrm{d}a_{12} a_{12} M^{abl}_{[ij]} \left(\sqrt{a_{12}^2 + z^2} \right) P_l \left(\frac{z}{\sqrt{a_{12}^2 + z^2}} \right) \frac{128\pi^5}{c_a c_b} \tag{4.44}$$

In order to have a reference for the strength of the interaction potentials we can

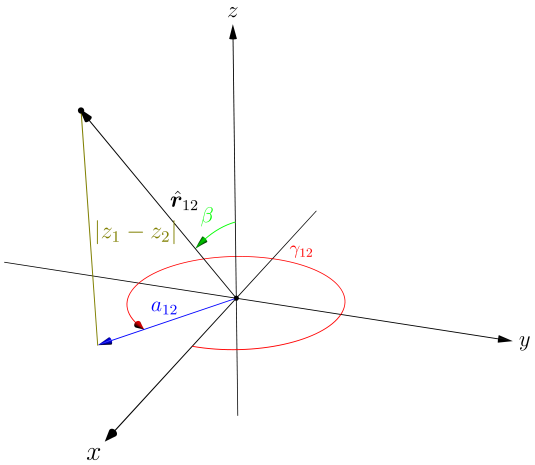


Figure 4.2: Description of the coordinates used in the partial integration of the mean-field kernels.

take a look at the corresponding moments of the mean-field potential that result

form the random-phase approximation. They take the following form

$$V^{mnl}_{[ij]}(r) = \epsilon^{mnl}_{[ij]}\phi_{\mathrm{sw}}(r) \tag{4.45}$$

$$\epsilon^{mnl}_{[ij]} = \frac{2l+1}{256\pi^5}\int \mathrm{d}\Omega_1\,\mathrm{d}\Omega_2\,\mathrm{d}\hat{\mathbf{r}}_{12}\overline{\Phi^{mnl}_{[ij]}}\sum_{\alpha,\beta=1}^{N_\mathrm{p}}\phi_\mathrm{p}(\mathbf{r}_{12},\hat{\mathbf{r}}_1^\alpha(\Omega_1),\hat{\mathbf{r}}_2^\beta(\Omega_2))\,. \tag{4.46}$$

The integral in Equation (4.46) cannot be solved analytically, which means that we have to resort to numerical methods.

To evaluate the Euler–Lagrange equation it is necessary to compute the partial derivatives of the excess free energy functional, with respect to both the density and the orientation distribution. They are given by

$$\frac{\delta\mathcal{F}^{\mathrm{mf}}_{\mathrm{ex,or}}[n,\alpha]}{\delta n(z)} = \sum_{i,j}\alpha^i(z)\left[n\alpha^j * M^{ij}\right] + \alpha^j(z)\left[n\alpha^i * \overline{M}^{ij}\right] \tag{4.47}$$

$$\frac{\delta\mathcal{F}^{\mathrm{mf}}_{\mathrm{ex,or}}[n,\alpha]}{\delta\alpha^k(z)} = \sum_{i,j} n(z)\left[n\alpha^j * M^{kj}\right] + n(z)\left[n\alpha^i * \overline{M}^{ik}\right] \tag{4.48}$$

It turns out that the contribution of the mean-field excess free energy cancels with partially with the entropic ODF contribution in the EL equation of the density profile. The entropic contribution in Equation (4.13) can be written out as

$$\begin{aligned}
-\langle \ln K\alpha\rangle_\alpha(\mathbf{r}) &= -\int \mathrm{d}\Omega\,\alpha(z,\Omega)\log(K\alpha(z,\Omega)) \\
&= \log\frac{1}{K}\int \mathrm{d}\Omega\psi(z,\Omega) + \int \mathrm{d}\Omega\,\alpha(z,\Omega)\sum_{lmn}\frac{2l+1}{K}\overline{D^l_{mn}}(\Omega)\frac{1}{n(z)}\frac{\delta\beta\mathcal{F}^{\mathrm{mf}}_{\mathrm{ex,or}}[n,\alpha]}{\delta\alpha^l_{mn}} \\
&= \log\frac{1}{K}\int \mathrm{d}\Omega\psi(z,\Omega) + \sum_{lmn}\frac{\alpha^l_{mn}(z)}{n(z)}\frac{\delta\beta\mathcal{F}^{\mathrm{mf}}_{\mathrm{ex,or}}[n,\alpha]}{\delta\alpha^l_{mn}}
\end{aligned} \tag{4.49}$$

where in the second line we inserted Equation (4.20) for the ODF α. In the case of mean-field ansatz the following equality holds

$$\sum_{lmn}\frac{\alpha^l_{mn}(z)}{n(z)}\frac{\delta\beta\mathcal{F}^{\mathrm{mf}}_{\mathrm{ex,or}}[n,\alpha]}{\delta\alpha^l_{mn}} = \frac{\delta\beta\mathcal{F}^{\mathrm{mf}}_{\mathrm{ex,or}}[n,\alpha]}{\delta n} \tag{4.50}$$

cf. Equations (4.47) and (4.48). This means that the EL equation for the density has no contribution from the mean-field free energy, as long as there are only interactions between M^{ij} with $i, j \neq 0$.

4.4.1 Random phase approximation

Before optimizing the interaction kernels using machine learning methods, it is useful to see first how well the random phase approximation handles the orientational correlation. We know already that the quality of RPA for the density profiles is rather weak, resulting from the small patch–patch bonding volume (see Section 3.1), which averages over orientation. This time we take the orientational degrees of freedom into account. For this, we need to solve the integral Equation (4.46), which we do, using Monte Carlo integration. The integral

$$I = \int \mathrm{d}\Omega_1 \, \mathrm{d}\Omega_1 \, \mathrm{d}\mathbf{r}_{12} f(\mathbf{r}_{12}, \Omega_1, \Omega_2) \tag{4.51}$$

is approximated by

$$I \approx \frac{1}{N} \sum_{i=1}^{N} f(\mathbf{r}_{12}^{(i)}, \Omega_1^{(i)}, \Omega_2^{(i)}) \tag{4.52}$$

where $x^{(i)}$ is a randomly sampled value of x. We can partially test the integration routine by comparing to the exact result in Equation (3.11). Indeed, using only 10^6 points for the MC integration we get a value of 3458.0 ± 11.6, while the exact value is

$$V_{00}^{000} = 256\pi^2\chi^2 = 3454.8 \tag{4.53}$$

For the higher moments there is unfortunately no way to check the correctness, however, we can check whether the correct symmetry conditions, derived in Section 2.6, are fulfilled. To this end we show the first moments in Figure 4.3, and higher ones in Figure 4.4 We see that both the correct matrix element are non-vanishing and that their relative strength also follows the predictions of Table 2.2. Finally, in order to find the RPA potential, the moments $V_{[ij]}^{mnl}(r)$ need to be converted from radial to cylindrical coordinates. The forward and backward transformations are given in Appendix A. The corresponding integrals need to be solved numerically.

The self-consistent results are shown in Figure 4.6. Similar to the isotropic density, the orientational profiles predicted by RPA are too weak (by more than one order of magnitude). Further there is barely any temperature dependence visible, though the functional form of the orientation profiles is not too far off.

4.4.2 Finding optimal parameters using ML

The general approach used in the training of most machine learning methods is gradient based. Let the model M depend on parameters θ, which need to be

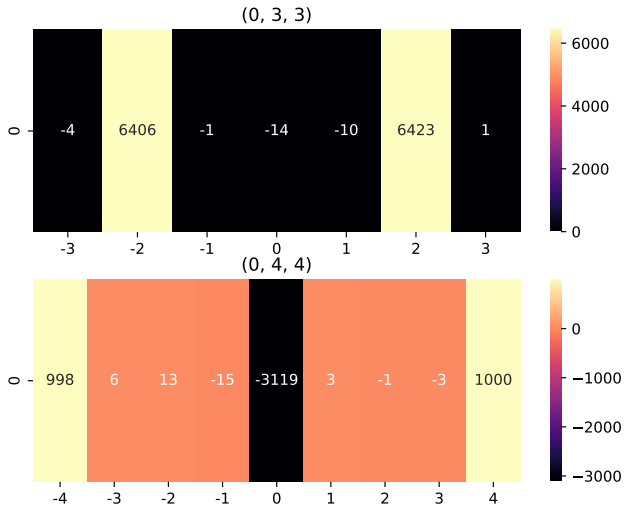


Figure 4.3: Individual RPA potentials computed using Monte Carlo integration and subsequent conversion to cylindrical symmetry.

optimized. Optimality is gauged by the value of the loss function ℓ for a certain output of the model $M(\theta) = m_\theta$ in comparison to ground truth data $\overline{m}$

$$\ell(\theta) = d(m_\theta, \overline{m}) \tag{4.54}$$

We want to find the parameters $\hat{\theta}$ that result in the smallest loss value, with the complication that the dimensionality of θ is quite high. Automatic differentiation and back-propagation allow us to efficiently evaluate the gradient of ℓ, therefore one can follow this gradient (in the negative direction) to further decrease the loss

$$\theta_{n+1} = \theta_n - \alpha \nabla \ell(\theta_n) \tag{4.55}$$

This is the basic idea behind the optimization procedure that we use. One important condition for this to work, is that all transformations happening in the model $M(\theta)$ need to be differentiable w.r.t. the variable θ. Our network is implemented using the Python library JAX [Bra+18], which offers an interface modelled after the popular NumPy library. Additionally, one is able to take gradients of most expressions, even control structures like for–loops. Through the use of just-in-time (JIT) compilation and execution on graphical processing units (GPUs) we can attain fast execution times with relatively little effort.

The actual iterative step of our model then consists of

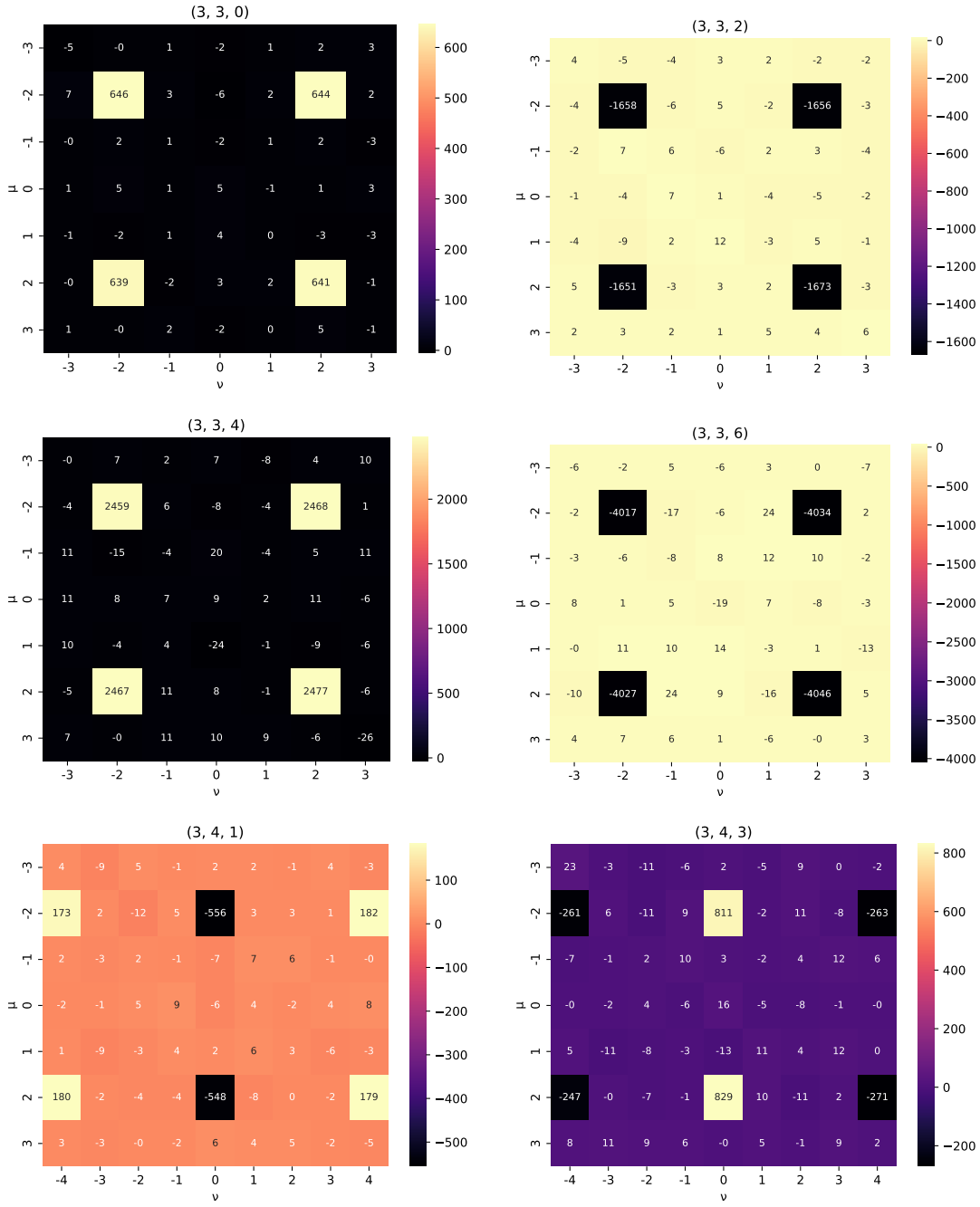


Figure 4.4: The Kern–Frenkel potential projected onto the spherical invariants $\Phi_{\mu\nu}^{mnl}$. The title shows the index tuple (m, n, l) corresponding to the respective matrix $V_{\mu\nu}^{mnl}$. Note that due to selection rules of the $3j$-symbol, the sum $m + n + l$ must be even.

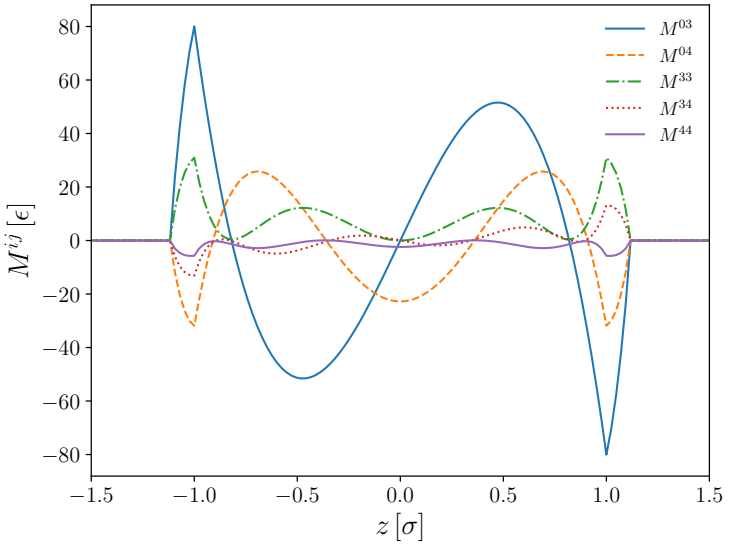


Figure 4.5: Individual RPA potentials computed using Monte Carlo integration and subsequent conversion to cylindrical symmetry.

1. Evaluating the non-normalized orientation distribution by

$$\psi(z,\Omega;\theta) = \exp\left(-\frac{1}{n(z)}\sum_{lmj}\frac{\delta\beta\mathcal{F}_{\mathrm{ex}}[n,\alpha;\theta]}{\delta\alpha^{l}_{m[j]}(z)}\overline{\Delta^{l}_{m[j]}}(\Omega)\frac{2l+1}{K}\right) \tag{4.56}$$

2. Normalizing

$$\alpha(z,\Omega;\theta) = \psi(z,\Omega;\theta)/\int \mathrm{d}\Omega'\psi(z,\Omega';\theta) \tag{4.57}$$

3. Projecting out the moments of interest

$$\alpha^{l}_{m[j]}(z;\theta) = \frac{2l+1}{K}\int \mathrm{d}\Omega\alpha(z,\Omega;\theta)\overline{\Delta^{l}_{m[j]}}(\Omega) \tag{4.58}$$

Up to this point we described the generative step of the network. That is, given a set of parameters θ, we have a prescription of how to compute the resulting predictions, in this case the density or orientation profile. However, the situation is slightly more complicated because what we are really interested in, is conditioning on the fixed point of the parametrized free energy functional and not just the output at some point. Taking the exact Euler–Lagrange equation into account we can theoretically get around the problem of solving for the fixed point since we know that it must be equal to the equilibrium distribution and those are given by simulations. When simulation data is too noisy, the evaluation at just one point

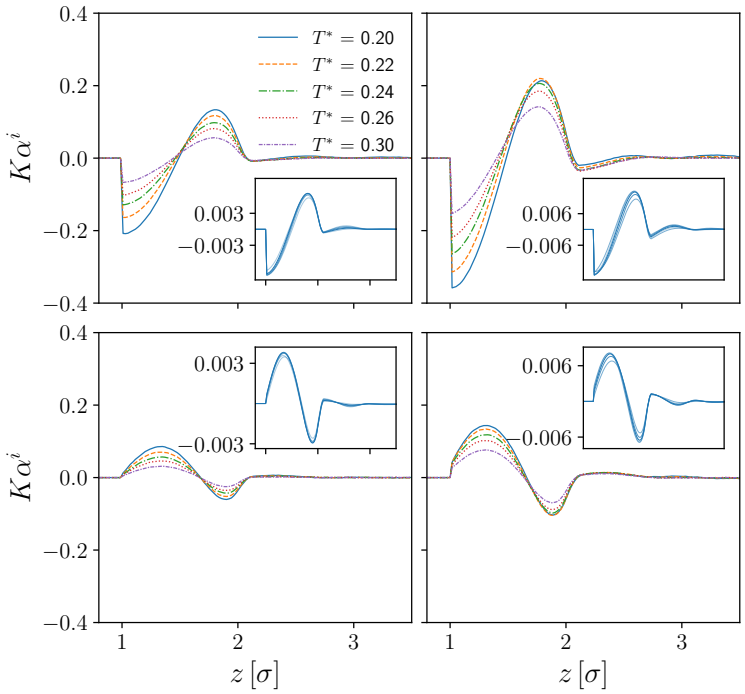


Figure 4.6: Orientational moments at a hard wall from simulations and self-consistent results using the mean-field free energy with the RPA kernels, shown in the insets. One can see that the RPA results severely underestimates the orientations. Further the temperature dependence of the RPA results is also not comparable to the simulation results, which show a strong dependence.

will become problematic as it pushes the network to learn false features stemming from statistical fluctuations.

To mitigate this, we propose to skip the evaluation at the known approximate fixed point and to solve the fixed point equation self-consistently in every forward pass. This, on the other hand, also becomes computationally impractical without further approximations as one would very soon hit memory limitations when computing the gradients of the loss functions using back-propagation.

We use the so-called *phantom gradient* to get around this problem, which is based on the implicit function theorem. Let $g(x, \theta)$ be the generative step of the EL-equation where we include the density profile and the orientation distribution in the variable x. We are interested in the fixed point $x^\star$ of g, given the parameters $\theta^\star$, which is equivalent to

$$g(x^\star, \theta^\star) - x^\star = 0 \tag{4.59}$$

By defining a function f with

$$f(x, \theta) = g(x, \theta) - x \tag{4.60}$$

we reformulate the original FP condition to $f(x^\star, \theta^\star) = 0$. In the following we will need the implicit function theorem, which states that for $f : \mathbb{R}^{m+n} \to \mathbb{R}^m$, $x \in \mathbb{R}^m$, $\theta \in \mathbb{R}^n$ and given that $f(x^\star, \theta^\star) = 0$ and the Jacobian

$$J_{f,\theta}(x^\star, \theta^\star) \equiv \left[\frac{\partial f_i}{\partial \theta_j}(x^\star, \theta^\star) \right] \tag{4.61}$$

is invertible, then there exists an open set $U \subset \mathbb{R}^n$ containing $\theta^\star$ and a function[1] $x^\star(\theta)$, such that

$$x^\star(\theta^\star) = x^\star \tag{4.62}$$

$$f(x^\star(\theta), \theta) = 0 \quad \forall \theta \in U \tag{4.63}$$

Moreover, we can compute the Jacobian of $x^\star(\theta)$ given only the Jacobian of f by

$$\left[\frac{\partial x_i^\star}{\partial \theta_j}(\theta) \right]_{m \times n} = - \left[J_{f,x}(x^\star(\theta), \theta) \right]^{-1}_{m \times m} \left[J_{f,\theta}(x^\star(\theta), \theta) \right]_{m \times n} \tag{4.64}$$

Noting that for the Jacobians we have

$$J_{f,x}(x^\star(\theta), \theta) = J_{g,x}(x^\star(\theta), \theta) - \mathbf{1} \tag{4.65}$$

$$J_{f,\theta}(x^\star(\theta), \theta) = J_{g,\theta}(x^\star(\theta), \theta) \tag{4.66}$$

[1] Note that $x^\star(\theta)$ is now written with an argument, highlighting the fact that it is a mapping.

we can write the derivative of the fixed point $x^\star$ w.r.t. the parameters θ as

$$\frac{\partial x^\star}{\partial \theta} = \left(\frac{\partial g}{\partial \theta}\right)_{x^\star,\theta} \left[\mathbf{1} - \left(\frac{\partial g}{\partial x}\right)_{x^\star,\theta}\right]^{-1} \tag{4.67}$$

which is the quantity we were originally interested in. The matrix inverse on the left-hand side still poses a problem as it needs to be numerically solved, which can be done using iterative methods. If the inversion is too costly or becomes numerically unstable one can expand the inverse in a Neuman series, which for bounded linear operators T says that

$$(\mathbf{1} - T)^{-1} = \sum_{k=0}^{\infty} T^k \tag{4.68}$$

Truncating after the zeroth order leaves us with

$$\frac{\partial x^\star}{\partial \theta} \approx \left(\frac{\partial g}{\partial \theta}\right)_{x^\star} \tag{4.69}$$

which tells us that we can compute the derivative of the fixed point at θ by determining the fixed point first and then evaluating the derivative of the fixed point equation itself at the point θ. In practice, this means that we can proceed by finding the fixed point for a certain parametrization without building up the computational graph, as we do not need to do a back-propagation on it. Therefore, memory consumption will be kept manageable and the only difficulty that might arise is the time needed to find the fixed point and whether one exists at all. To speed up the convergence we use Anderson acceleration instead of the simpler Picard iteration scheme. It works by keeping previous iterations outputs and builds up the next trial input by finding the optimal linear combination of the previous ones, similar to DIIS. In fact, it can be shown that both approaches are mathematically equivalent. After having found the fixed point $x^\star$ one only needs to evaluate the Euler–Lagrange function g once, this time including also the computational graph, as we need to back-propagate on the variable θ.

Compared to the evaluation on the known equilibrium solution alone we need to put in more effort, but we gained an important advantage. Namely, that the parametrization does in fact result in a fixed point at all, because it is numerically stable. What it is not clear yet, is over which range it is stable. For every data point in the training set, we begin the fixed point iteration with a uniform density profile of the same bulk density as in the training set. Bulk densities that exceed the ones from the training set are not encountered, and we cannot be certain that

the fixed point equation remains stable at those densities. One might therefore argue that it is beneficial to include into the training range also values of higher density than one is originally interested in (probably on physical grounds, as high density profiles mostly get dominated by the hard-core interaction), as then the network is automatically regularized over the whole density range.

The loss is defined as follows

$$\ell = \mathrm{MSE}\left(\alpha^{\star}, \alpha_{\mathrm{eq}}\right) + \lambda \sum_{ij} L^1(M^{ij}) \tag{4.70}$$

where the second term is the L^1 (absolute value) norm of the parametrized interaction potentials, and the corresponding hyperparameter λ that controls how strongly large parameter values are penalized. This regularization, together with the self-consistency of the solutions, allows us to systematically include physically relevant parts into the mean-field kernels.

4.4.3 ML results for the mean-field

In the present ansatz the only temperature-dependent term is the factor β that makes the free energy derivatives dimension-less, see Equations (4.19) and (4.20). The mean-field ansatz does not have an explicit temperature dependence. Since we know however, that the free energy is a temperature-dependent quantity we solve this issue by learning the kernels for every temperature in the training set separately. This amounts to having different mean-field potentials $M^{ij}(T_k^*)$ for every temperature T_k^* in the training set.

No density interaction For reference, we repeat the EL equation for density distribution in the anisotropic framework, which reads

$$n(\mathbf{r}) = \exp\left(-\langle \ln K\alpha\rangle_{\alpha}(\mathbf{r}) + \beta\mu - \beta V^{\mathrm{ext}}(\mathbf{r}) - \frac{\delta\beta\mathcal{F}_{\mathrm{ex}}[n,\alpha]}{\delta n(\mathbf{r})}\right) \tag{4.71}$$

As described earlier, the angular averaged entropic term $-\langle \ln K\alpha\rangle_{\alpha}(\mathbf{r})$ cancels the functional derivative of the excess free energy

$$-\frac{\delta\beta\mathcal{F}_{\mathrm{ex}}[n,\alpha]}{\delta n(\mathbf{r})} \tag{4.72}$$

as long as there are only mean-field components M^{ij} with $i, j > 0$, see also Equation (4.49). The remaining entropic term

$$\log \frac{1}{K} \int \mathrm{d}\Omega \psi(z, \Omega) \tag{4.73}$$

is fixed by the orientational distribution and therefore not adjustable. Its effect can be computed from the "empiric" value of α determined from simulation and is rather small. Nonetheless, this ansatz is attractive from a physical point of view. Since every term in the mean-field excess free energy contains at least one orientational moment, which vanish in the bulk phase, we are not changing the underlying equation of state, which is already approximated rather well by the Stopper–Wu functional.

In Figure 4.7 we show the self-consistent results for the symmetrized orientational moments α^3 and α^4 after the mean-field kernels have been learned. The accuracy is better for the high temperature state point across all densities in the training set, which is expected from a mean-field ansatz. At low temperatures, apart from a worse agreement, we also observe a stronger variation at different densities, with lower densities being less accurately described.

With density interaction We train a second ansatz with density corrections by allowing mean-field terms of the form M^{00}, i.e. terms that do not vanish in the bulk and are not canceled by the entropic term. Apart from changing the density profiles, this also modifies the equation of state. The resulting self-consistent moments after training are shown in Figure 4.8. We observe similar results as with the training without the density correcting terms, although the profiles are slightly more accurate. This is due to the fact that the density profiles that enter the calculation of α^3 and α^4 also change. The new "corrected" density profiles are shown in Figure 4.9. The trend from earlier carries over with the high temperatures state points being reproduced accurately over the whole density range and the low temperature state points slightly worse. Different from the orientational moments, this time the low density profiles are more accurate, because here the difference between simulation and the StWu reference functional are quite small to begin with. Furthermore, we see that for the highest densities unphysical density modulations start to emerge with increasing density.

Kernels An example of the resulting kernels with density correction (for the best hyperparameter set) is shown in Figures 4.10 and 4.11. Compared to the RPA kernels, we can see that the magnitude is much larger, which is expected as the

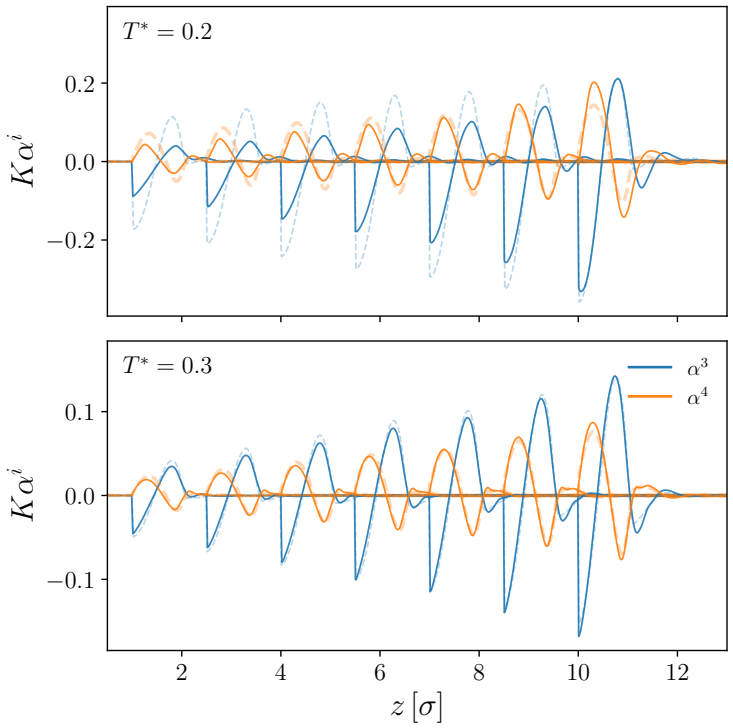


Figure 4.7: Self-consistent solution for the orientational moments in the mean-field ansatz *without* the density correction term. The results for several bulk densities are shown at once, by starting with the lowest density profile on the l.h.s. and plotting profiles with increasing density with a z-axis offset. We see very good agreement for the high temperature case and qualitative agreement at low temperatures, where the magnitude of the orientation is too low for all but the highest density profile, where it overshoots the data.

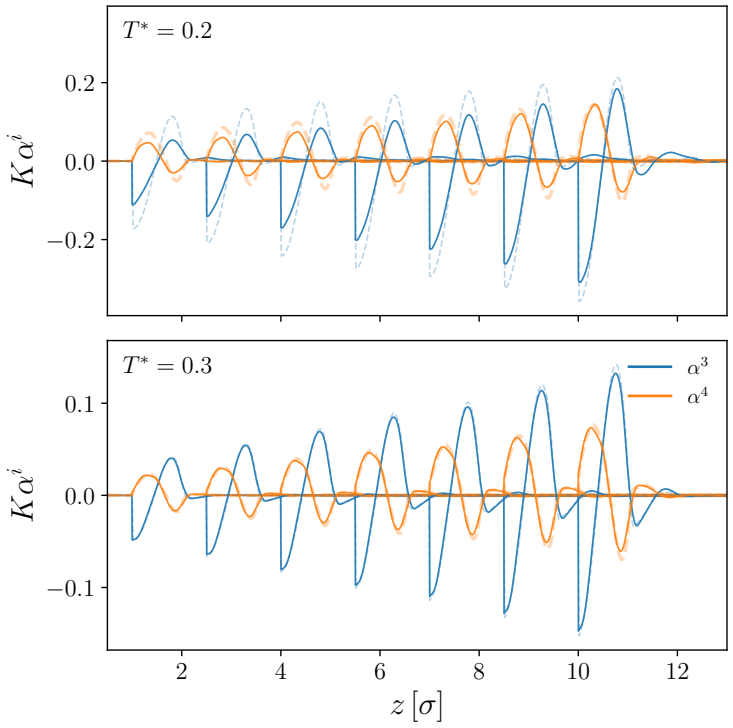


Figure 4.8: Self-consistent solution for the orientational moments in the mean-field ansatz *with* the density correction term and same density encoding as in Figure 4.7. Similarly, the high-temperature regime agrees well with simulations. With respect to the low-temperature regime we see an improvement, with all profiles still slightly smaller than the simulation data.

RPA severely underestimates the orientational correlation. The relative strength of the individual potentials roughly agrees with that of the RPA results, meaning that potentials with lower indices are larger, especially at higher temperatures.

The loss function in Equation (4.70) includes a regularizer λ that penalizes large (absolute) values of the interaction potentials M^{ij}. It is customary in machine learning application to introduce a regularizer into the loss function to prevent the network from overfitting. In our special case we expect the interaction potentials to be rather short-ranged, although the ideal range is not known. Here the use of a regularizer can help to restrict the support of the learned kernels. In Figure 4.12 we show the resulting mean-field potentials for varying parameter λ. The strength of the regularizer λ has the expected effect on the kernels. One can nicely see how interactions that lead to the larger moments emerge first and smaller ones

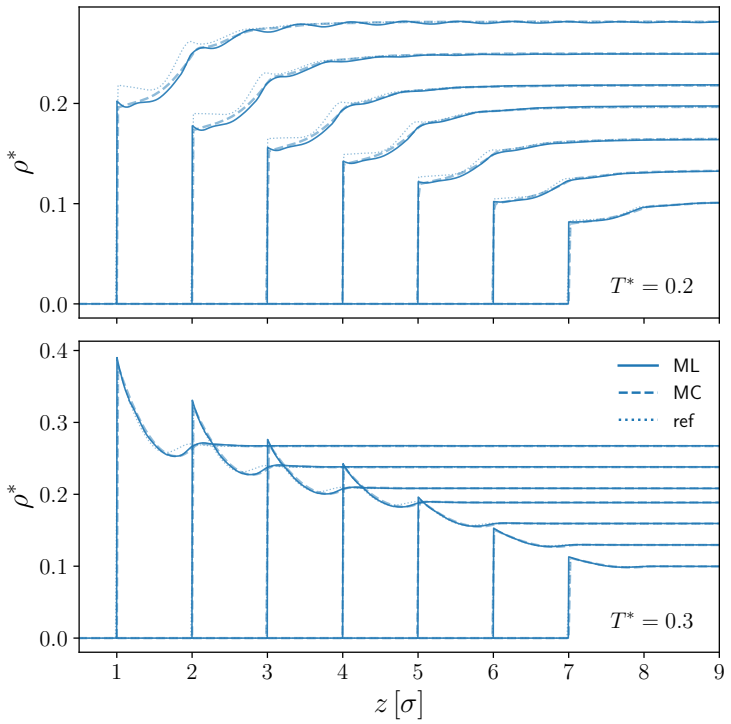


Figure 4.9: Self-consistent density profiles for the mean-field ansatz with density correction (labelled by "ML"). The lines labelled as "ref" correspond to the Stopper–Wu functional. We see that there is good agreement between the simulation data and the trained model, for most state points. For low temperature and high densities however, oscillations start to become visible.

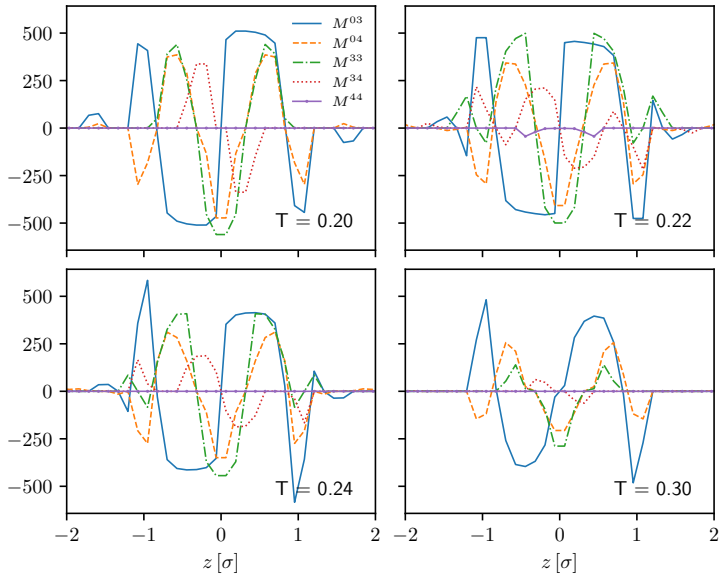


Figure 4.10: Isothermal interaction potentials (mean-field kernels) M^{ij} with at least one orientational moment involved, for the ansatz with density correction.

following later on.

Self-consistency As discussed earlier, the self-consistency (SC) of the learned functional parametrization is in theory guaranteed by employing the phantom gradient when updating the parameters of the network. This allows us to solve the EL-equations even if the mean squared error in Equation (4.70) is still comparatively large, which can either happen early on in training or when the regularizer λ is sufficiently strong.

In an ansatz that doesn't require self-consistency, i.e. by evaluating the EL equations only at the fixed point, this would not be possible, since in that case the network parameters would be determined by minimizing the residual

$$g(x^\star, \theta) - x^\star = \Delta \tag{4.74}$$

for a given fixed point (equilibrium value) $x^\star$. One can only expect to get converging results of the iteration scheme once the value of Δ has reached a certain threshold. Adding a regularizer term to the loss, like we did in Equation (4.70), will lead to competition between the regularizer term and the residual Δ, and it might happen that the threshold for self-consistency is never reached.

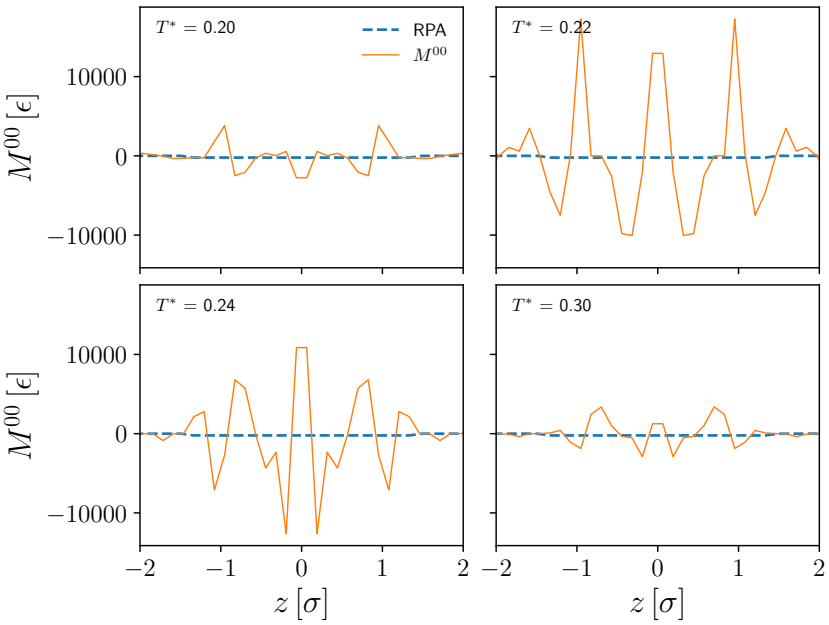


Figure 4.11: Mean-field kernel for the density–density interaction, i.e. the density correction case. The corresponding RPA potential is barely visible as it is much smaller. Further, one observes that the temperature-dependence is not monotonous.

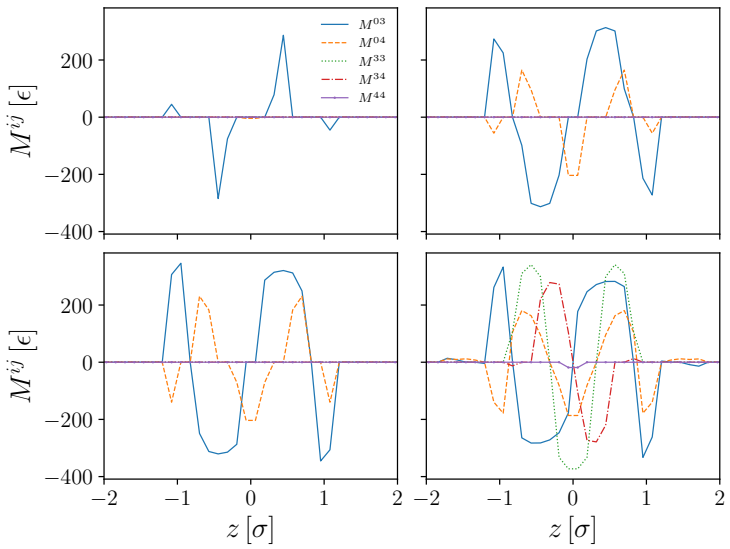


Figure 4.12: Mean-field kernels with decreasing value of regularizer $\lambda \in \left[5 \cdot 10^{-3}, 5 \cdot 10^{-4}, 1 \cdot 10^{-4}, 1 \cdot 10^{-5}\right]$ in reading direction. One can see, that lowest order interactions start to emerge first.

We can test our approach by solving the fixed point equation for parameters that were trained using comparatively large values of the parameter λ. In this case the network is not able to accurately fit the observed orientation and density profiles. Consequently, the MSE loss in Equation (4.70) is rather high. Nonetheless, we are able to solve the EL equation self-consistently as shown in Figure 4.13

Extrapolation capabilities One of our original aims was to find a functional that also works for state points that lie outside the training data set. Usually, extrapolation performance is measured using a validation data, i.e. keeping a separate dataset which is not used during training. However, in our case we chose to maximize the amount of data that goes into training, since it needs to be able to perform at different state points (with regard to the density) and the amount of data is quite limited. The formalism then allows us to relatively easily asses the out-of-training-set performance later on, by tuning the chemical potential μ. We used the learned mean-field potentials to compute the equilibrium values for two higher density states, namely $\rho_b = 0.344, 0.391$ (corresponding to average densities in the box of 0.35 and 0.4, respectively) at the highest and lowest temperature in our training sets. Note that we do not "extrapolate" in the temperature regime as this would not work since the MF potentials are learned for every temperature

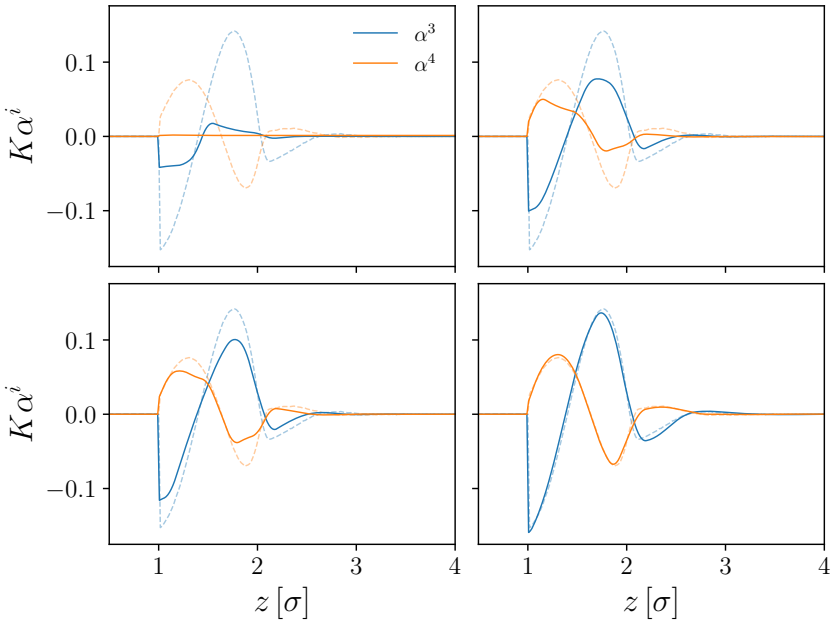


Figure 4.13: Self-consistent results of the learned mean-field potentials (shown in Figure 4.12) for different values of the regularized λ. For high values of λ the potentials are held as small as possible, thus producing profiles which do not agree with simulations perfectly. Nonetheless, we are able to solve the EL equations for every value of λ.

separately because no temperature-dependence was included in our ansatz.

The corresponding results are shown in Figure 4.14. We observe that the high-temperature state point is predicted very accurately. The low-temperature state point however shows again the appearance of unphysical oscillations, similar to Figure 4.9. When further increasing the density, the fixed point iteration becomes numerically unstable and no solutions can be found.

4.5 Tuning the Wertheim bonding term

We have seen before that the Stopper–Wu functional actually works reasonably well, the discrepancies notwithstanding. These can chiefly be summarized as peak positions that are slightly off for high temperatures, and while critical drying is present at low temperatures, a hard-sphere structure is still too dominant, with peaks that are too pronounced.

One can pose the question of whether it is possible to correct these deficiencies by modifying the weight functions that enter the bonding term of the free energy. In fact, it is not clear why fundamental measure weights should be used at all, when we are treating bond–bond interactions which are not directly related to the geometry of the sphere.

We proceed by introducing a new set of weighted densities m_ν, that enter only in the bonding part of the free energy

$$m_\nu = \rho \otimes w'_\nu = \rho \otimes (w_\nu + \varpi_\nu) \tag{4.75}$$

which becomes

$$\mathcal{F}_{\text{ex}}[\rho] = \mathcal{F}_{\text{hs}}[\{n_\nu\}] + \mathcal{F}_{\text{bond}}[\{m_\nu(\theta)\}] \tag{4.76}$$

The kernel corrections ϖ are added to the original ones from FMT and could in principle cancel them out completely, there is hence complete freedom for the form of the kernels. We keep the sum structure to facilitate training, such that we can start at a stable point. Also note that by learning all m_ν independently we break the dependence between certain kernels that is present in FMT[2].

The network is trained similar as before, such that in every iteration we solve for the fixed point and then use the phantom gradient to evaluate the approximate gradient. The external potential being now a function of r, we can realize a potential

[2] E.g. w_0 is related to w_2.

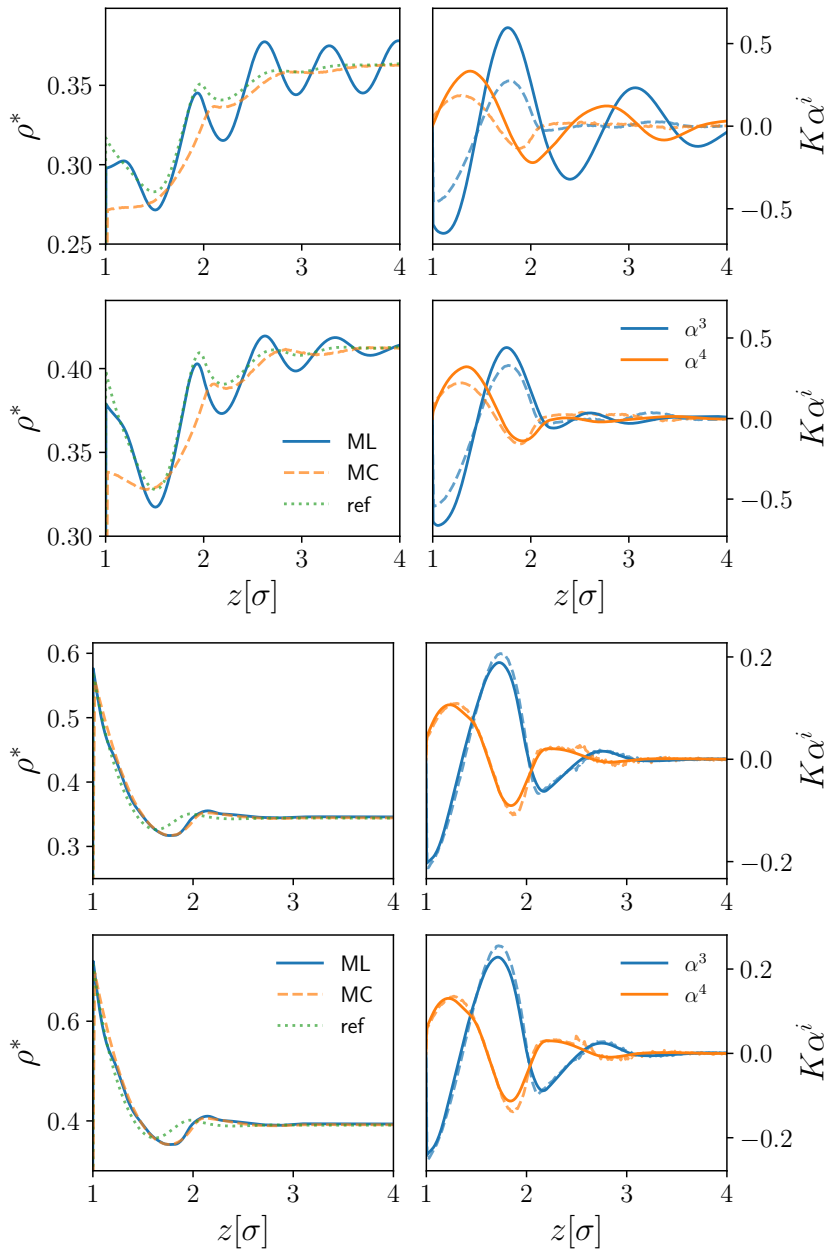


Figure 4.14: Self–consistent solutions of the learned MF potential at densities higher than the training set (average densities of $\rho^* = 0.35$ and 0.4, respectively.. The density profiles of the low temperature isotherm ($T^* = 0.20$, upper plot) show clear oscillations, that are not yet present in the high temperature case ($T^* = 0.30$, lower plot). The oscillations are unphysical and make the iteration scheme numerically unstable.

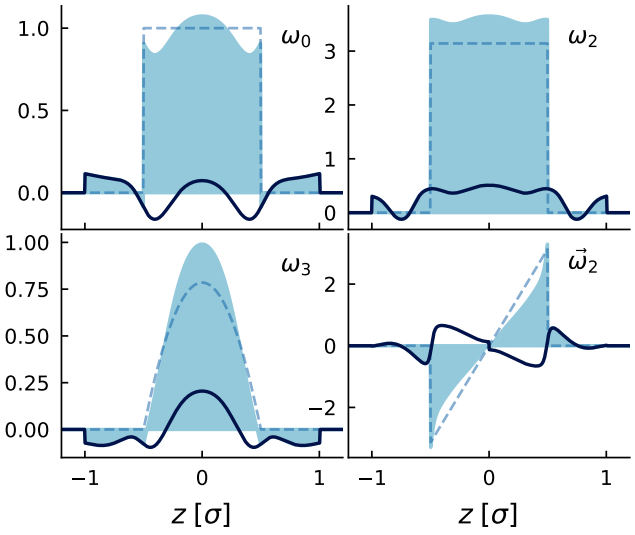


Figure 4.15: The improved kernels in real space. The dashed lines show the original FMT weights, while the thick lines represent the learned modifications ϖ_ν. The resulting kernels are shown as filled areas.

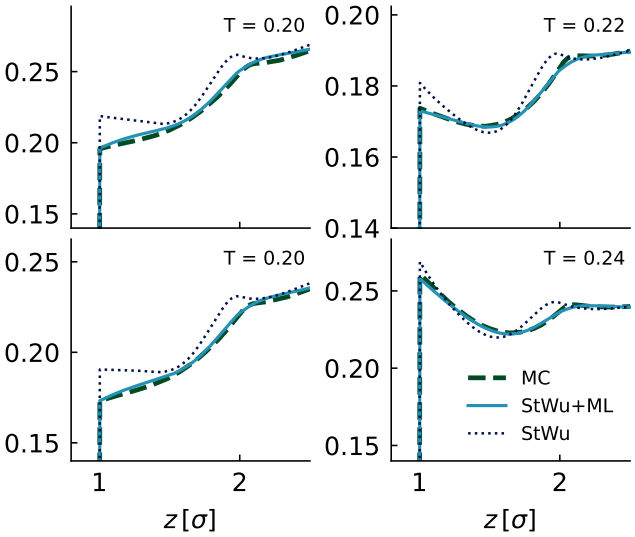


Figure 4.16: Density profiles resulting from the new set of optimized weights. We observe very good agreement with simulation data at all temperatures and densities.

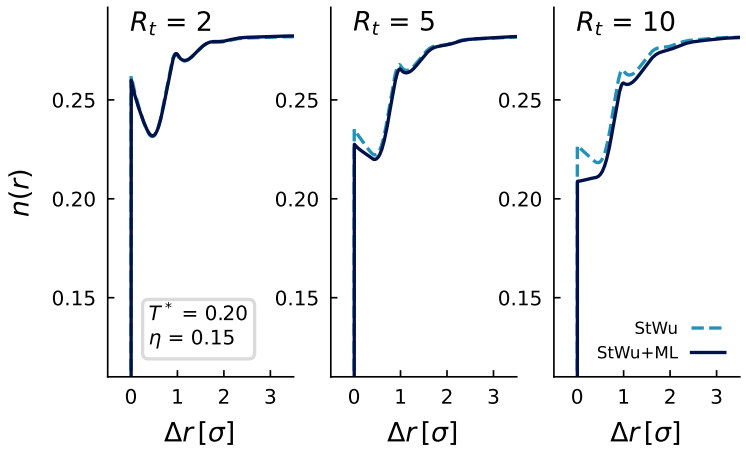


Figure 4.17: Density profiles around a hard spherical tracer of variable radius R_t. In the limit $R_t \to \infty$, the hard-wall behavior is recovered, which can be observed in the right-most plot. For small tracer radii however, the improved kernels don't modify the profiles noticeably.

that corresponds to the potential exerted by a spherical particle with particle center at $r = 0$.

The kernels that were learned using wall-geometry data can also be used for radial problems. This is in contrast to the anisotropic mean-field ansatz from above, where different quantum numbers l were mixed, making this inversion impossible. We add a cutoff for the maximal width of the corrected kernels at twice the width of the original kernels. The exact value is a trade-off between stability of training (long range kernels introduce instabilities) and the accuracy of the new kernels (a wider support allows in general for more tunable degrees of freedom). No further regularization is added to the loss.

Results In Figure 4.15 we show the new set of improved kernels together with the reference FMT kernels. Due to the self-consistent training procedure we observe again smooth kernels in real space (up to the cutoff at $z_{\max} = \pm 1$).

Furthermore, we show in Figure 4.16 the main result, namely the self-consistent density profiles resulting from the new kernels. Overall we see very good agreement between the StWu+ML profiles and the MC profiles. Note, that in contrast to the mean-field ansatz, where every temperature was learned separately, here we learn the same kernels for all temperatures. The temperature dependence is same as in the "normal" Stopper–Wu functional.

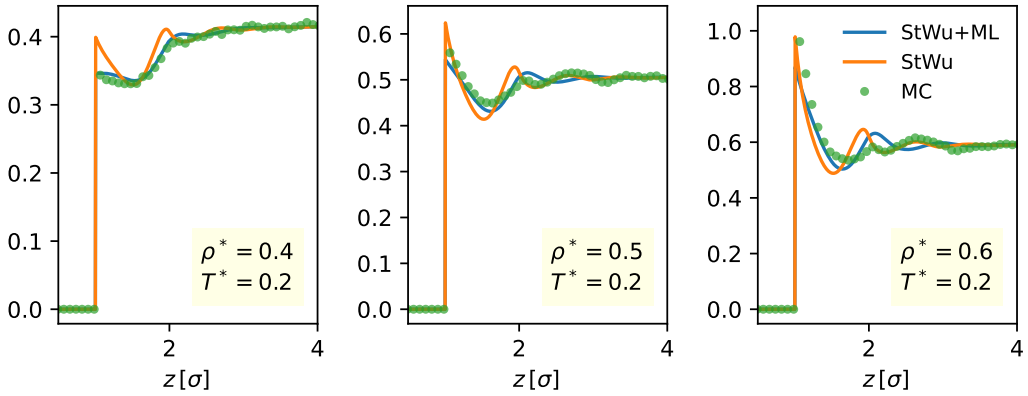


Figure 4.18: Equilibrium density profiles for the original StWu and the StWu+ML functional evaluated at high densities, outside the training region. The dotted data points are the results from simulation. Generally, better agreement is observed with the StWu+ML functional, although at high densities the structure changes drastically and is not well captured by either functional.

Compared to the MF-ansatz the numerical stability outside the training region is increased. Even at bulk densities that are twice as high as the largest value in the training set, the EL can be solved. We show a selection of these extrapolated curves in Figure 4.18. While the agreement between StWu+ML and the simulation data is worse than in the training set, it is nevertheless better than the results from the standard StWu functional except for very high density state points, where the true structure is very different due to packing effects.

Direct correlation function We can compute the resulting direct correlation function from the functional by taking the derivative of the free energy density w.r.t. the density twice. The contribution separate into the part from the HS functional, plus the one from the bonding term (Stopper–Wu), now with modified kernels w_i'

$$c_{\mathrm{iso}}(r) = -\sum_{ij} \left.\frac{\partial^2 \phi_{\mathrm{bond}}}{\partial m_i\, \partial m_j}\right|_{\rho_\mathrm{b}} \left(\pm w_i' \otimes w_j'\right)(r) + c_{\mathrm{hs}}(r) \tag{4.77}$$

Because the four weight functions that enter the bonding term are learned individually, the previous dependence between the original FMT kernels w_0 and w_2 no

longer holds. This has in particular an effect on the calculation of the DCF, where it is necessary to compute the convolutions between the individual kernels, as seen in Equation (4.77). By using the fact that the original vector-valued kernel $\mathbf{w}_2$ could be written as a gradient of the scalar-valued kernel, the convolution could be reduced to a scalar convolution, which is not the case for the learned kernels. More information about how to perform these convolutions can be found in Appendix A. The resulting DCFs are shown in Figure 4.19 for different values of T^* and ρ^*. For reference, we also show the integral equations results (Percus–Yevick closure) for the hard sphere and the sticky hard sphere. Let us first compare the "exact" results (c^{000}) with those from the standard StWu functional. Since the StWu functional is based on the FMT weighted densities, the resulting DCF can only be non-zero in the range $r \in [0, 1\sigma]$, because further out the convolutions of the weight kernels vanish. This means that the "bonding correlation" in the range $r \in [1\sigma, 1.119\sigma]$ cannot be reproduced accurately. Instead, we observe a correlation that is rising steadily and finally goes to zero discontinuously at $r = 1\sigma$. The correlation inside the overlap region $r < 1\sigma$ is overestimated when compared to the simulation result.

The results of the ML-improved StWu functional show a different behavior. Due to the extended range of the weight functions, we now start to see a positive correlation in the bonding region. Also, the repulsive part $r < 1\sigma$ is much closer to the simulations results.

Seeing that the DCF of the ML-improved functional performs closer to simulation is a valuable consistency check. Especially since the network was not trained on this particular metric, but only the density profiles.

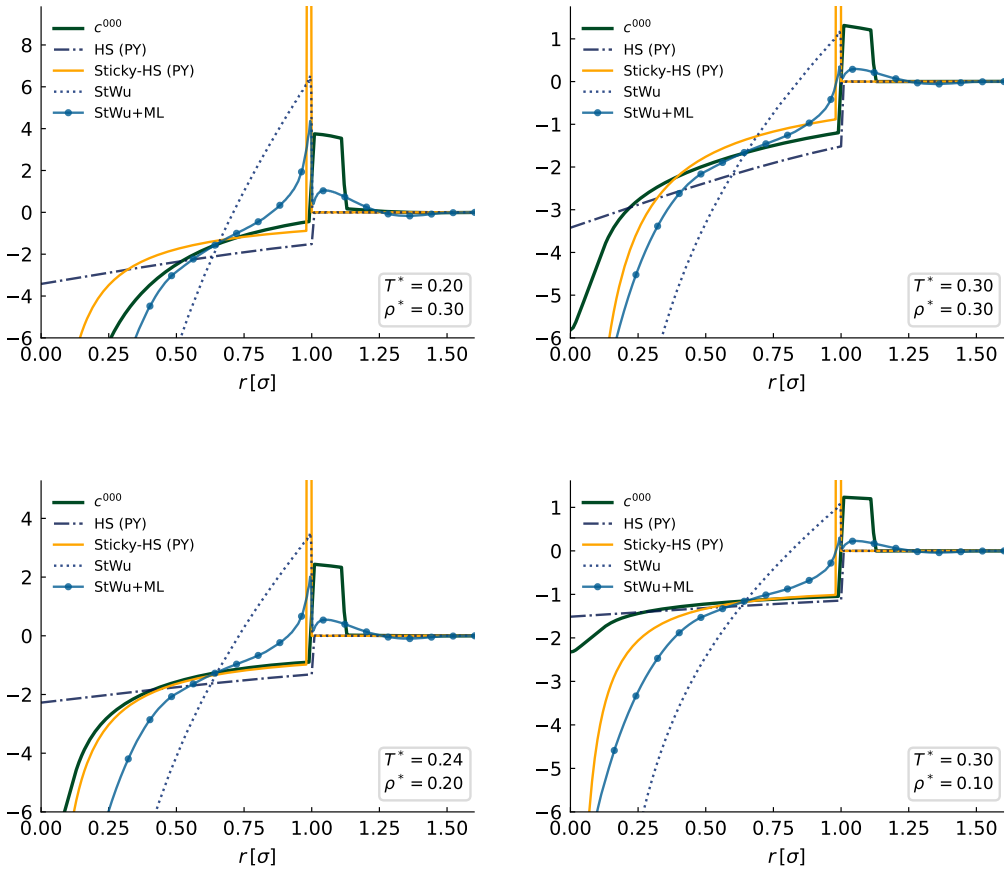


Figure 4.19: Comparison between direct correlation functions from MC simulations (c^{000}, angular averaged), hard sphere Percus–Yevick (PY), sticky hard spheres (PY), the Stopper–Wu functional (StWu) and the ML-augmented Stopper–Wu functional (StWu+ML) at different state points (shown in plot).

5 Conclusion and outlook

In this thesis, we explored the potential of using machine learning to develop a density functional capable of accurately describing the structure of anisotropic associating liquids. Our investigation began with examining the orientational structure of a Kern–Frenkel liquid at a hard wall and around spherical tracer particles. Given the strong orientational ordering observed, we proposed enhancing previous DFT approaches by incorporating orientational information into the functional. Specifically, we introduced a machine-learned mean-field ansatz that effectively predicts orientational moments and, to a lesser extent, density profiles.

A second approach leverages the degrees of freedom available in the family of FMT functionals, namely the weighted density kernels. The fixed functional form of the functional improves the numerical stability of the learned functional, resulting in highly accurate density profiles. However, this raises questions about whether a complete description, encompassing both density and orientation distribution, is necessary or if achieving precise density profiles suffices.

The self-consistency of the learned mean-field kernels emerged as a crucial factor, ensured through implicit differentiation. Difficulties arise with respect to the numerical stability of the trained mean-field functional. For bulk densities that are far from the training region, the Euler–Lagrange equations cannot be solved reliably. Learning improved (weighted density) kernels, while leaving the functional fixed, seems to results in a larger region of stability. In this regard it would be worthwhile to look into methods that improve the stability of the learned functional outside the training region.

One way this could be achieved is through the use of more diverse training data for example in the form of an external potential $V(z, \Omega)$ that acquires a dependency with respect to the orientational configuration Ω of the particles. This kind of "anchoring potential" has been considered before in [CGS16; WRS18].

A further point worth investigating is the use of more elaborate ansätze for the excess free energy. Our experiments with the equation learner network (EQL) for anisotropic liquids, building on the approach pioneered in [LMO20], showed

suboptimal results. The method, termed functional equation learner (FEQL), is based on representing the desired quantity (here the excess free energy functional) as a feed-forward network with nonlinear activation functions, serving as basis functions of the expansion (for more information, see also [ML16; SLM18]). The big advantage of this network is that it learns symbolic approximation of the mapping, which can be read of directly and interpreted by a human. The network starts in a randomly initialized stat and is trained on its output. Sparsity (i.e. simplicity) of the final equation is achieved by adding a regularization term to the loss, which penalizes large absolute values of the weights inside the linear layers. Finally, all weights below a certain threshold are set to zero. An alternative regularization method is that of a stochastic L_0 loss, see [MMT17; LWK18], which only minimizes the number of non-vanishing weights, but has no effect on their magnitude.

While the EQL architecture is quite successful in finding symbolic input–output mappings for general datasets, it is less well suited to our problem. The reason for this is the additional indirection that enters through the iterative solution of the EL equations. With the random initialization we cannot make sure that there is always a self-consistent solution to the FP equations, hence it is not possible to back-propagate through the network. At the same time, starting with very small or vanishing weights is also not very beneficial, as then the self-consistent solutions correspond to the trivial isotropic orientation distribution, again not delivering any helpful gradient information. A possible workaround could be to start from a parametrization that allows for self-consistent solutions of the EL equations that at the same time must not be zero. A different method, which is also able to learn symbolic representations of the unknown excess free energy functional was used in the conceptually similar formalism of quantum density functional theory, see [Ma+22]. It was shown that genetic algorithms combined with reinforcement learning are able to find state-of-the-art approximations in an automated way.

Other machine learning approaches to enhance DFT calculations focus on neural approximations for $c^{(1)}(\mathbf{r})$ or $F_{\text{ex}}[\rho;\theta]$. The former [Sam+23; SS24] involves solving the principal Euler–Lagrange equation [cf. Equation (2.86)] for $c^{(1)}$, which can be determined from simulations, given the density profile and external potential. However, this is not feasible for anisotropic particle density distributions, because we cannot solve the EL equation for the anisotropic DCF.

The latter approach [Dij+24] learns a neural approximation of the excess free energy functional using homogeneous radial distribution functions from simulations which are mapped onto corresponding $c^{(2)}(\mathbf{r})$ correlation functions, by solving the Ornstein–Zernike equation. The parametrized excess free energy $F_{\text{ex}}[\rho;\theta]$ is differentiated twice (using automatic differentiation) and the weights optimized

such that the correlation function of the parametrized functional reproduces that from the simulation/OZ route. While the solution of $c^{(2)}$ through the OZ equation is theoretically also possible for anisotropic two-body potentials [Sim+25], it incurs significant computational costs. For a more detailed exposition of machine learning methods related to DFT, see [SO24].

Appendix A

Convolutions in different geometries

When evaluating the Euler–Lagrange equation for the equilibrium density distributions we often have to evaluate convolutions, e.g. for the weighted densities

$$n(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \rho(\mathbf{r}') w(\mathbf{r} - \mathbf{r}') = \rho \otimes w \tag{A.1}$$

Similarly, when evaluating the DCF for FMT functionals we need to evaluate convolutions between the weight funtions themselves. The convolution theorem states that

$$\rho \otimes w = \mathscr{F}^{-1}\left[\left[\mathscr{F}[\rho]\,\mathscr{F}[w]\right]\right] \tag{A.2}$$

where $\mathscr{F}, \mathscr{F}^{-1}$ denote the Fourier transform and its inverse. In cartesian coordinates the convolution separates into three individual convolutions, one for every dimension, which can be readily performed by the Fast Fourier transform. When working on problems with planar symmetry the density profile $\rho(z)$ will only depend on the z-coordinate and the integral in Equation (A.1) reduces to

$$n(z) = \int \mathrm{d}x' \,\mathrm{d}y' \,\mathrm{d}z' \rho(z') w(\mathbf{r} - \mathbf{r}') \tag{A.3}$$

meaning that the integration over x and y can be performed without knowledge of ρ. The integral $\int \mathrm{d}x' \,\mathrm{d}y' w(\mathbf{r}')$ corresponds to a projection from radial geometry $w(\mathbf{r})$ to planar geometry, resulting in $w(z)$. Similar conversion need to be applied when we learn convolution kernels in planar geometry, but want to use them in radially symmetric problems.

For a general, radially symmetric function we define by $\mathscr{P}[f(r)](z)$ the projection onto the z-axis, and by $\mathscr{P}^{-1}[f]$ the inverse operation. Note that the function can

also be vector-valued. The projections are calculated as follows

$$\mathscr{P}[f(r)](z) = \int_{\mathbb{R}^2} \mathrm{d}x\,\mathrm{d}y f\left(\sqrt{x^2+y^2+z^2}\right) \tag{A.4}$$

$$= \int_0^{2\pi} \mathrm{d}\gamma \int_0^{\infty} \mathrm{d}a\, a f\left(\sqrt{a^2+z^2}\right) \tag{A.5}$$

$$= 2\pi \int_z^{\infty} \mathrm{d}u\, u f(u) \tag{A.6}$$

where we used the substitution $u = \sqrt{a^2+z^2}$. For vector-valued functions we assume again that $\mathbf{f}(r)$ is radially symmetric, i.e. $\mathbf{f}(r) = \mathbf{e}_r f(r)$. The unit radial vector at a point $\mathbf{r} = (x, y, z)$ is written in cartesian coordinates

$$\mathbf{e}_r = \frac{1}{r}\left(x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z\right) \tag{A.7}$$

Integrating $\mathbf{e}_r$ over the x- and y-axis will make the results vanish as the positive and negative contributions cancel. We therefore only need to consider the z-component

$$\mathscr{P}[\mathbf{f}(r)](z) = \int_{\mathbb{R}^2} \mathrm{d}x\,\mathrm{d}y\, z\mathbf{e}_z f\,(r)\,/r \tag{A.8}$$

$$= z\mathbf{e}_z \int_0^{2\pi} \mathrm{d}\gamma \int_0^{\infty} \mathrm{d}a\, a f\left(\sqrt{a^2+z^2}\right)/\left(a^2+z^2\right) \tag{A.9}$$

$$= 2\pi z\mathbf{e}_z \int_z^{\infty} \mathrm{d}u\, u f(u)/u \tag{A.10}$$

$$= 2\pi z\mathbf{e}_z \int_z^{\infty} \mathrm{d}u f(u) \tag{A.11}$$

With these we can calculate the projected weight functions. They are

$$\mathscr{P}[w_2(r)](z) = 2\pi \int_z^{\infty} \mathrm{d}u\, u\delta(R-|u|) \tag{A.12}$$

$$= 2\pi R\Theta(R-|z|) \tag{A.13}$$

where the Θ-function is a result of the integration limits ($|z|$ must be smaller than R for the integral in Equation (A.12) to be non-zero). Further we have

$$\mathscr{P}[w_3(r)](z) = 2\pi \int_z^{\infty} \mathrm{d}u\, u\Theta(R-|u|) \tag{A.14}$$

$$= 2\pi \int_z^{R} \mathrm{d}u\, u\Theta(R-|z|) \tag{A.15}$$

$$= \pi\left(R^2-z^2\right)\Theta(R-|z|) \tag{A.16}$$

$$\mathscr{P}[\mathbf{w}_2(r)](z) = 2\pi z \mathbf{e}_z \int_z^\infty \mathrm{d}u \delta(R - |u|) \tag{A.17}$$

$$= 2\pi z \mathbf{e}_z \Theta(R - |z|) \tag{A.18}$$

The other weights are again just scalar multiples of the previous three. The inverse operation (going from planar to radial) are given by

$$\mathscr{P}^{-1}[f(z)](r) = -\frac{1}{2\pi r} \left.\frac{\mathrm{d}F(z)}{\mathrm{d}z}\right|_{z=r} \tag{A.19}$$

$$\mathscr{P}^{-1}[\mathbf{f}(z)](r) = -\frac{1}{2\pi}\mathbf{e}_r \left.\frac{\mathrm{d}(F(z)/z)}{\mathrm{d}z}\right|_{z=r} \tag{A.20}$$

and can be checked by substitution.

Fourier transforms of radially symmetric function The three dimensional Fourier transform of a radially symmetric function $f(r)$, written in spherical coordinates reads

$$\mathscr{F}[f(r)](k) = \int_0^\infty \int_0^\pi \int_0^{2\pi} r^2 \sin\theta \,\mathrm{d}\phi \,\mathrm{d}\theta \,\mathrm{d}r f(r) e^{-2\pi i r k \cos\theta} \tag{A.21}$$

The angular integral can be performed by subsitution and using the complex exponential representation of the sine function, resulting in

$$\mathscr{F}[f(r)](k) = \frac{2}{k}\int_0^\infty \mathrm{d}r r f(r) \sin(2\pi r k) = \frac{2}{k}\mathscr{S}[rf(r)] \tag{A.22}$$

where $\mathscr{S}$ is the sine transform. The remaining transformations are given by

$$\mathscr{F}^{-1}[f(k)](r) = \frac{2}{k}\mathscr{S}[rf(r)] \tag{A.23}$$

$$\mathscr{F}[\mathbf{f}(r)](k) = \mathbf{e}_k \frac{2i}{k}\mathscr{C}[rf(r)] - \mathbf{e}_k \frac{i}{\pi k^2}\mathscr{S}[f(r)] \tag{A.24}$$

$$\mathscr{F}^{-1}[\mathbf{f}(k)](r) = \mathbf{e}_r \frac{i}{\pi r^2}\mathscr{S}[f(k)] - \mathbf{e}_r \frac{2i}{r}\mathscr{C}[f(k)k] \tag{A.25}$$

where for the vector-valued functions we also need the cosine transfrom $\mathscr{C}$. For more details on the numerical aspects of convolutions in the context of FMT, see [Sti+20].

Appendix B

Formulas

B.1 Expansions

Here, we collect the most important formulas concerning expansion in Wigner D-matrices and the rotationally invariant expansion.

Expansion in Wigner D-matrices

$$f(\mathbf{r},\Omega) = \sum_{lmn} f^l_{mn}(\mathbf{r}) D^l_{mn}(\Omega) \tag{B.1}$$

$$f^l_{mn}(\mathbf{r}) = \frac{2l+1}{K} \int \mathrm{d}\Omega f(\mathbf{r},\Omega) \overline{D^l_{mn}}(\Omega) \tag{B.2}$$

Expansion in invariant linear combinations

$$f(\mathbf{r},\Omega) = \sum_{imj} f^i_{m[j]}(\mathbf{r}) \Delta^i_{m[j]}(\Omega) \tag{B.3}$$

$$f^i_{m[j]}(\mathbf{r}) = \frac{2i+1}{K} \int \mathrm{d}\Omega f(\mathbf{r},\Omega) \overline{\Delta^i_{m[j]}}(\Omega) \tag{B.4}$$

$$\Delta^i_{m[j]}(\Omega) = \sum_n d^i_{[j]}(n) D^i_{mn}(\Omega) \tag{B.5}$$

Rotational invariant expansion

$$f(\mathbf{r}, \Omega_1, \Omega_2) = \sum_{mnl\mu\nu} f_{\mu\nu}^{mnl}(r_{12}) \Phi_{\mu\nu}^{mnl}(\Omega_1, \Omega_2, \hat{\mathbf{r}}) \tag{B.6}$$

$$f_{\mu\nu}^{mnl} = (2l+1) \int \mathrm{d}\Omega_1 \, \mathrm{d}\Omega_2 \, \mathrm{d}\hat{\mathbf{r}} f_{\mu\nu}^{mnl}(\hat{\mathbf{r}}, \Omega_1, \Omega_2) \overline{\Phi_{\mu\nu}^{mnl}}(\Omega_1, \Omega_2, \hat{\mathbf{r}}) \tag{B.7}$$

$$\Phi_{\mu\nu}^{mnl}(\Omega_1, \Omega_2, \hat{\mathbf{r}}) = c_m c_n \sum_{\mu'\nu'\lambda'} \begin{pmatrix} m & n & l \\ \mu' & \nu' & \lambda' \end{pmatrix} \overline{D_{\mu'\mu}^{m}}(\Omega_1) \overline{D_{\nu'\nu}^{n}}(\Omega_2) \overline{D_{\lambda'0}^{l}}(\hat{\mathbf{r}}) \tag{B.8}$$

$$f_{\mu\nu\chi}^{mn}(r) = \sum_{l} \begin{pmatrix} m & n & l \\ \chi & -\chi & 0 \end{pmatrix} f_{\mu\nu}^{mnl}(r) \tag{B.9}$$

$$f_{\mu\nu}^{mnl}(r) = (2l+1) \sum_{l} \begin{pmatrix} m & n & l \\ \chi & -\chi & 0 \end{pmatrix} f_{\mu\nu\chi}^{mn}(r) \tag{B.10}$$

$$f_{\mu\nu\chi}^{mn}(r) = \int \mathrm{d}\omega_1 \, \mathrm{d}\omega_2 f(\omega_1, \omega_2, r) D_{\chi\mu}^{m}(\omega_1) D_{\underline{\chi}\nu}^{n}(\omega_2) \tag{B.11}$$

B.2 Functional derivatives

From the defining equation of the expansion we can deduce the functional derivative w.r.t. to the ODF $\alpha(\mathbf{r}, \Omega)$ is written in terms of the orientational moments. It follows that

$$\begin{aligned} \frac{\delta}{\delta\alpha(\mathbf{r}, \Omega)} &= \int \mathrm{d}\mathbf{r}' \sum_{lmn} \frac{\delta}{\delta\alpha_{mn}^{l}(\mathbf{r}')} \frac{\delta\alpha_{mn}^{l}(\mathbf{r}')}{\delta\alpha(\mathbf{r}, \Omega)} \\ &= \int \mathrm{d}\mathbf{r}' \sum_{lmn} \frac{\delta}{\delta\alpha_{mn}^{l}(\mathbf{r}')} \frac{2l+1}{K} \delta(\mathbf{r} - \mathbf{r}') \overline{D_{mn}^{l}}(\Omega) \\ &= \sum_{lmn} \frac{\delta}{\delta\alpha_{mn}^{l}(\mathbf{r})} \frac{2l+1}{K} \overline{D_{mn}^{l}}(\Omega) \end{aligned} \tag{B.12}$$

If we chose to expand the ODF in the invariant linear combinations Δ instead of the unmodified Wigner D-matrices, we have structurally the same result, as the Δ

do not mix Wigner D-matrices of different l. In this case we get

$$\begin{aligned}\frac{\delta}{\delta\alpha(\mathbf{r},\Omega)} &= \int \mathrm{d}\mathbf{r}' \sum_{ij} \frac{\delta}{\delta\alpha^i_{(j)}(\mathbf{r}')} \frac{\delta\alpha^i_{(j)}(\mathbf{r}')}{\delta\alpha(\mathbf{r},\Omega)} \\ &= \int \mathrm{d}\mathbf{r}' \sum_{ij} \frac{\delta}{\delta\alpha^i_{(j)}(\mathbf{r}')} \frac{2i+1}{K} \delta(\mathbf{r}-\mathbf{r}') \overline{\Delta^i_{(j)}}(\Omega) \\ &= \sum_{ij} \frac{\delta}{\delta\alpha^i_{(j)}(\mathbf{r})} \frac{2i+1}{K} \overline{\Delta^i_{(j)}}(\Omega)\end{aligned} \tag{B.13}$$

Appendix C

Additional data

C.1 Symmetrized Wigner D-matrices

In Table 2.2 we show the tetrahedrally invariant Wigner D-matrices for $l \leq 11$. They were determined using the Python package `SymPy` which is able to handle symbolic mathematical manipulations. Unfortunately for higher values of l the computations take an unreasonable amount of time, and the coefficients would need to be converted to floating point numbers in order to get results. Here we limit ourselves to symbolic results only.

C.2 Invariant expansion of the KF potential

Here we show higher invariant expansion moments of the Kern–Frenkel potential in order to check the validity of the basis functions listed in Table 2.2.

basis function	components
$\Delta^3_{m[1]}$	$+\frac{\sqrt{2}}{2}D^3_{m\underline{2}} + \frac{\sqrt{2}}{2}D^3_{m2}$
$\Delta^4_{m[1]}$	$+\frac{\sqrt{30}}{12}D^4_{m\underline{4}} - \frac{\sqrt{21}}{6}D^4_{m0} + \frac{\sqrt{30}}{12}D^4_{m4}$
$\Delta^6_{m[1]}$	$+\frac{\sqrt{7}}{4}D^6_{m\underline{4}} + \frac{\sqrt{2}}{4}D^6_{m0} + \frac{\sqrt{7}}{4}D^6_{m4}$
$\Delta^6_{m[2]}$	$-\frac{\sqrt{10}}{8}D^6_{m\underline{6}} - \frac{\sqrt{22}}{8}D^6_{m\underline{2}} + \frac{\sqrt{22}}{8}D^6_{m2} + \frac{\sqrt{10}}{8}D^6_{m6}$
$\Delta^7_{m[1]}$	$+\frac{\sqrt{33}}{12}D^7_{m\underline{6}} - \frac{\sqrt{39}}{12}D^7_{m\underline{2}} - \frac{\sqrt{39}}{12}D^7_{m2} + \frac{\sqrt{33}}{12}D^7_{m6}$
$\Delta^8_{m[1]}$	$+\frac{\sqrt{390}}{48}D^8_{m\underline{8}} - \frac{\sqrt{42}}{24}D^8_{m\underline{4}} + \frac{\sqrt{33}}{8}D^8_{m0} - \frac{\sqrt{42}}{24}D^8_{m4} + \frac{\sqrt{390}}{48}D^8_{m8}$
$\Delta^9_{m[1]}$	$+\frac{\sqrt{26}}{8}D^9_{m\underline{6}} + \frac{\sqrt{6}}{8}D^9_{m\underline{2}} + \frac{\sqrt{6}}{8}D^9_{m2} + \frac{\sqrt{26}}{8}D^9_{m6}$
$\Delta^9_{m[2]}$	$-\frac{\sqrt{21}}{12}D^9_{m\underline{8}} - \frac{\sqrt{51}}{12}D^9_{m\underline{4}} + \frac{\sqrt{51}}{12}D^9_{m4} + \frac{\sqrt{21}}{12}D^9_{m8}$
$\Delta^{10}_{m[1]}$	$+\frac{\sqrt{561}}{48}D^{10}_{m\underline{8}} - \frac{\sqrt{11}}{8}D^{10}_{m\underline{4}} - \frac{\sqrt{390}}{48}D^{10}_{m0} - \frac{\sqrt{11}}{8}D^{10}_{m4} + \frac{\sqrt{561}}{48}D^{10}_{m8}$
$\Delta^{10}_{m[2]}$	$-\frac{\sqrt{170}}{32}D^{10}_{m\underline{10}} - \frac{\sqrt{114}}{96}D^{10}_{m\underline{6}} + \frac{\sqrt{741}}{48}D^{10}_{m\underline{2}} - \frac{\sqrt{741}}{48}D^{10}_{m2} + \frac{\sqrt{114}}{96}D^{10}_{m6} + \frac{\sqrt{170}}{32}D^{10}_{m10}$
$\Delta^{11}_{m[1]}$	$+\frac{\sqrt{399}}{48}D^{11}_{m\underline{10}} - \frac{3\sqrt{3}}{16}D^{11}_{m\underline{6}} + \frac{\sqrt{510}}{48}D^{11}_{m\underline{2}} + \frac{\sqrt{510}}{48}D^{11}_{m2} - \frac{3\sqrt{3}}{16}D^{11}_{m6} + \frac{\sqrt{399}}{48}D^{11}_{m10}$

Table C.1: Linear combinations of Wigner matrices that are invariant under all tetrahedral group transformations up to $l = 11$

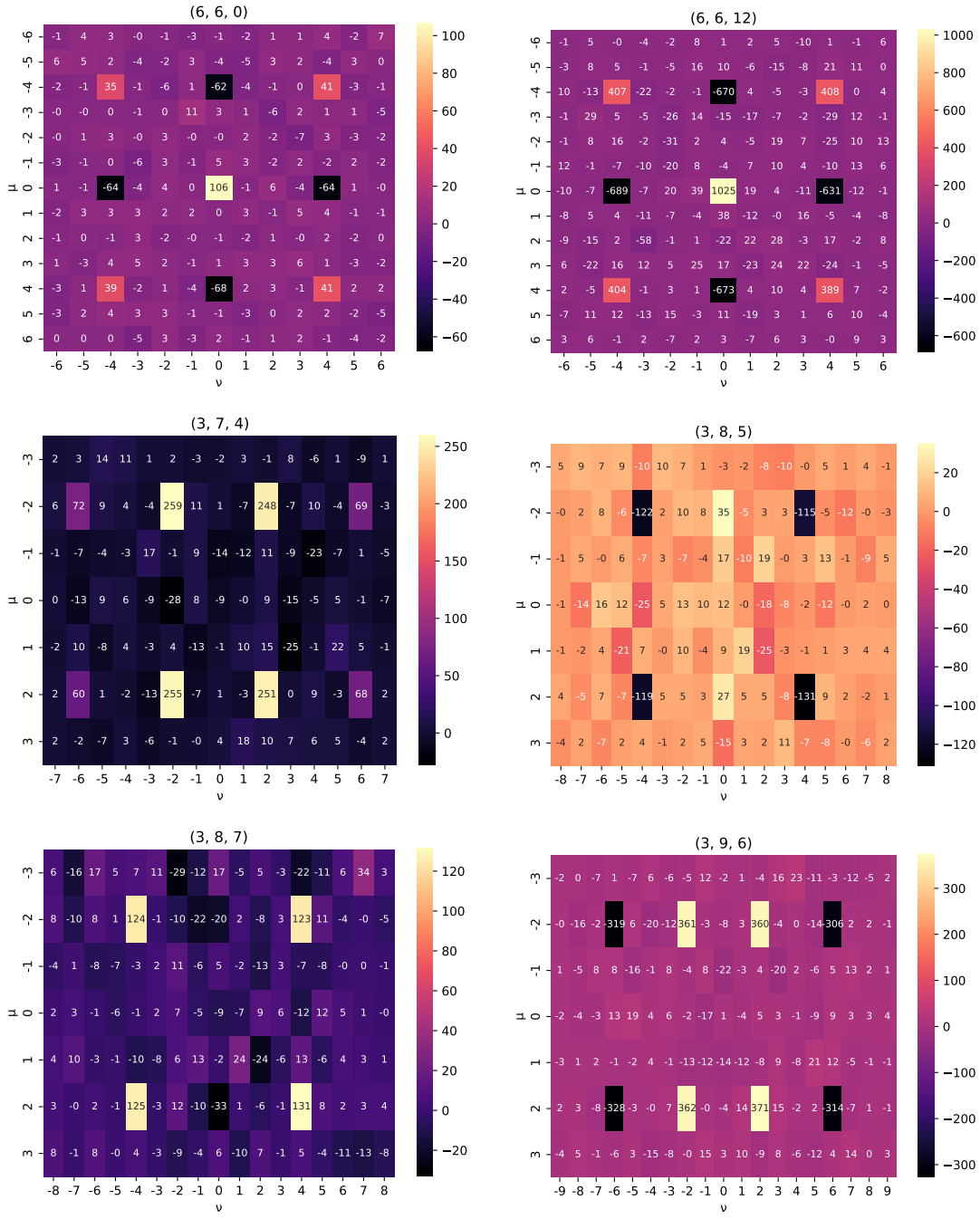


Figure C.1: The Kern–Frenkel potential projected onto the spherical invariants $\Phi_{\mu\nu}^{mnl}$. The title shows the index tuple (m, n, l) corresponding to the respective matrix $V_{\mu\nu}^{mnl}$. Note that due to selection rules the sum $m + n + l$ must be even.

Appendix D

Code listings

D.1 Recursive method

Here we demonstrate the implementation of the recursive method of evaluating the Wigner D-matrices after [Cho+99].

The algorithm is implemented in SymPy and returns a given matrix element D^l_{mn} evaluated at an angle triplet parametrized as rotation matrix R_{ij} (named X in the code). Further we show how it can be converted to C code making it easy to include it in existing code bases. The code to generate the necessary instructions and also test the implementation is given in Listing D.1

Listing D.1: Python code for the evaluation of Wigner D-matrices according to the method of Choi et al.

```
import sympy as sy
import numpy as np
import sys

# define coefficients, see Choi et al., p. 5
def a(l,m,mp):
    if abs(m) == l:
        return 0.
    elif (abs(mp) == l):
        return 0.
    else:
        return sy.sqrt(((l+m)*(l-m))/((l+mp)*(l-mp)))

def b(l,m,mp):
    if (m == -l or m == -l+1) :
        return 0.
    elif (abs(mp) == l):
        return 0.
    else:
        return sy.sqrt(((l+m)*(l+m-1))/(2.*(l+mp)*(l-mp)))


```

```
def c(l,m,mp):
    if (abs(m) == l or mp == -l or mp-1 == -l):
        return 0.
    else:
        return sy.sqrt((2.*(l+m)*(l-m))/((l+mp)*(l+mp-1)))


def d(l,m,mp):
    if (m == -l or m ==(-l+1) or mp == -l or mp-1 == -l):
        return 0.
    else:
        return sy.sqrt(((l+m)*(l+m-1))/((l+mp)*(l+mp-1)))


def _get_FG(R):
    sqrt2 = sy.sqrt(2)
    F = sy.Matrix([
        [(R[1, 1]+R[0,0])/2, R[0,2]/sqrt2, (R[1,1]-R[0,0])/2],
        [R[2,0]/sqrt2, R[2,2], -R[2,0]/sqrt2],
        [(R[1,1]-R[0,0])/2, -R[0,2]/sqrt2, (R[1,1]+R[0,0])/2]
    ])
    G = sy.Matrix([
        [(R[1, 0]-R[0,1])/2, R[1,2]/sqrt2, -(R[1,0]+R[0,1])/2],
        [-R[2,1]/sqrt2, 0, -R[2,1]/sqrt2],
        [(R[1,0]+R[0,1])/2, R[1,2]/sqrt2, (R[0,1]-R[1,0])/2]
    ])
    return F, G

def _get_D1():
    R = sy.MatrixSymbol("X", 3, 3)
    F, G = _get_FG(R)
    return F + sy.I * G

# generate C code for evaluating D^l_{mn}
def generate_code(l, m, n):
    def get_str(n):
        if n < 0:
            nstr = "m" + str(abs(n))
        else:
            nstr = str(n)
        return nstr

    m_str = get_str(m)
    n_str = get_str(n)

    fname = "d" + str(l) + m_str + n_str
    with open(fname+".gen", "w") as file:
        file.write(sy.ccode(Dlmn(l,m,n)))

    print(f"written {fname}.gen\n")


def get_random_rotation():
    from scipy.spatial.transform import Rotation as R
    rand_matrix = R.random().as_matrix()
    return rand_matrix

def get_matrix_fn(l,m,n):
    lamb = sy.lambdify(sy.MatrixSymbol("X", 3, 3), Dlmn(l, n, m), 'numpy')
    return lamb
```

```
83
84
85 def get_indices(max):
86         indices = []
87         for l in range(1, max+1):
88             for m in range(-l, l+1):
89                 for n in range(-l, l+1):
90                     indices.append((l, m, n))
91         return indices
92
93 # convert sympy expression to numpy function
94 def to_lambd(expr):
95     return sy.lambdify(sy.MatrixSymbol("X", 3, 3), expr, "numpy")
96
97
98 # reference implementation: ‘spherical‘ package
99 def eval_wigner(indices, rmatrix):
100     import spherical
101     import quaternionic
102
103     ell_max = 16
104     wigner = spherical.Wigner(ell_max)
105     result = []
106     for idx in indices:
107         result.append(wigner.D(quaternionic.array.from_rotation_matrix(rmatrix)
    )[wigner.Dindex(*idx)])
108     return np.array(result)
109
110
111 class get_D1:
112     def __init__(self):
113         d1 = _get_D1()
114         self.dmat = d1.T
115
116     def __getitem__(self, indices):
117         m = indices[0]+1
118         n = indices[1]+1
119         if m<0 or m>2:
120             return 0
121         if n<0 or n>2:
122             return 0
123         return self.dmat[m, n]
124
125 D1 = get_D1()
126
127 def Dlmn(l,m,n):
128     if n >= 0:
129         return ___Dlmn(l,m,n)
130     else:
131         return __Dlmn(l,m,n)
132
133 # recurrence A, see Choi et al, p. 5
134 def _Dlmn(l,m,n):
135     if l == 1:
136         return D1[m, n]
137     return a(l,m,n) * Dlmn(l-1,m,n) * D1[0,0] +\
138         b(l,m,n) * Dlmn(l-1,m-1,n) * D1[1, 0] +\
139         b(l,-m,n) * Dlmn(l-1,m+1,n) * D1[-1,0]
140
141 # recurrence C
```

```
def __Dlmn(l,m,n):
    if l == 1:
        return D1[m, n]
    return c(l,m,-n) * Dlmn(l-1,m,n+1) * D1[0,-1] +\
        d(l,m,-n) * Dlmn(l-1,m-1,n+1) * D1[1, -1] +\
        d(l,-m,-n) * Dlmn(l-1,m+1,n+1) * D1[-1,-1]

# recurrence D
def ___Dlmn(l,m,n):
    if l == 1:
        return D1[m, n]
    return c(l,m,n) * Dlmn(l-1,m,n-1) * D1[0,1] +\
        d(l,m,n) * Dlmn(l-1,m-1,n-1) * D1[1, 1] +\
        d(l,-m,n) * Dlmn(l-1,m+1,n-1) * D1[-1,1]

def test_implementation(lmax):
    rm = get_random_rotation() # perform comparison on random rotation

    indices = get_indices(lmax) # get valid indices
    refs = eval_wigner(indices, rm.T) # evaluate reference implementation

    for i, idx in enumerate(indices):
        impl = to_lambd(Dlmn(*idx))(rm) # eval recursive implementation

        if np.isclose(refs[i], impl):
            print(f"{idx} \t passed \t {refs[i]:.3f} {impl:.3f}")
        else:
            diff = np.abs(refs[i] - impl)
            print(f"{idx} \t failed \t difference = {diff}")


if __name__ == "__main__":
    # call test_implementation to test
    # otherwise generate code
    l = int(sys.argv[1])
    m = int(sys.argv[2])
    n = int(sys.argv[3])
    generate_code(l, m, n)
```

Evaluating the code for $l, m, n = 3, 0, 2$ by calling `python3 wignerd.py 3 0 2`, generated the file `d302.gen`, which can be called as follows

```
#include <math.h>
#include <complex.h>

double complex D302(const double* X){
    return
    #include "d302.gen"
    ;
}
```

D.2 Diagonalization method

The following code computes the Wigner D-matrix D^l_{mn}, column–wise for arbitrary l. The necessary eigenvectors (see main text) are calculated using LAPACK and the Eigen library. Note that the computations need to make use of complex matrices, while the resulting moments are real due to the tetrahedral symmetry.

Listing D.2: C++ code for evaluating Wigner D-matrix elements according to the method by Feng et al.

```
#include <iostream>
#include <complex>
#include <vector>
#include <memory>
#include <cassert>
#include <cstdio>

#define LAPACK_COMPLEX_CUSTOM
#define lapack_complex_double std::complex<double>
#define lapack_complex_float std::complex<float>
#include <lapacke.h>

#include <Eigen/Dense>


using Eigen::MatrixXcd;
using Eigen::VectorXcd;


class WignerD
{
public:
  WignerD(int lmax) : _lmax{lmax}
  {
    // cache eigenvector for later
    for (int l=0; l<=lmax; l++)
      {
  this->jEigenvecs.push_back(this->getEigenvecs(l));
      }
  }

  VectorXcd eval_col(int l, int n, double phi, double theta, double chi)
  {
    // evaluate one column of the wigner matrix. n is the physical index
    // i.e. going from -l to +l
    assert(l <= this->_lmax && "l is larger than lmax");
    assert(std::abs(n) <= l && "|m| must not be larger than l");

    // compute exp factor for phi and chi
    VectorXcd factor = (VectorXcd::LinSpaced(2*l + 1, -l, l)*phi*
      std::complex<double>(0., -1.0)).array().exp();
    factor *= std::exp(std::complex<double>(0., -1.)*chi*
           static_cast<double>(n));

    VectorXcd result(2*l + 1);
    const int limit = 2*l+1;
```

```
47      double mu;
48
49      for (int m=0; m<limit; m++)
50        {
51    std::complex<double> res = 0.0;
52    for (int i=0; i<limit; i++)
53      {
54        mu = static_cast<double>(i-l);
55        res += std::exp(mu* theta*std::complex<double>(0., -1.))*
56          (this->jEigenvecs[l](i,m))*std::conj(this->jEigenvecs[l](i,n+l));
57      }
58    result(m) = res;
59        }
60      return conj(result.array()*factor.array());
61    }
62
63
64  private:
65    MatrixXcd getEigenvecs(int j)
66    {
67      VectorXcd J = getJ(j).conjugate();
68      VectorXcd J0 = VectorXcd::Zero(2*j+1);
69
70      // matrix to diagonalize
71      MatrixXcd Jb(2*j+1, 2);
72      Jb << J0 , J;
73
74      const int n = 2*j+1; // size of the matrix
75      const int kd = 1; // number of super/sub-diagonals
76      const int ldab = 2*j+1;
77
78      // array for eigenvalues
79      double w[n];
80
81      // array for eigenvectors
82      MatrixXcd z(n, n);
83
84      // status info
85      int info;
86
87      info = LAPACKE_zhbev(LAPACK_ROW_MAJOR, 'V', 'L', n, kd, Jb.data(),
88         ldab, w, z.data(), n);
89
90      if (info == 0){
91  #ifdef DEBUG
92        std::cout << "Eigenvalues:" << std::endl;
93        for (int i = 0; i < n; i++) {
94    std::cout << w[i] << " ";
95        }
96        std::cout << std::endl;
97        std::cout << "Eigenvectors:" << std::endl;
98        std::cout << z << std::endl;
99  #endif
100     } else {
101       std::cerr << "Error occurred in zhbev. Info: " << info << std::endl;
102       exit(-1);
103     }
104     return z;
105   }
106
```

```
107   VectorXcd getJ(int j)
108   {
109     const VectorXcd idx = VectorXcd::LinSpaced(2*j + 1, 1-j, j + 1);
110     const auto x0 = idx.array() + j;
111     const auto x1 = (j+1) - idx.array();
112     VectorXcd result = -(x1*x0).sqrt()/std::complex(0., 2.0);
113     result(2*j) = 0.0;
114     return result;
115   }
116   int _lmax;
117   std::vector<MatrixXcd> jEigenvecs;
118 };
```

This revision: 8ae4159a0f44350922298acafb1e2ac521115ea1 on branch master.

Acknowledgements

Firstly, I wish to express my sincere thanks to Martin Oettel and Georg Martius for allowing me to dive into this fascinating project that exists at the delightful intersection of physics and machine learning. Their guidance was invaluable, and their trust in allowing me the freedom to explore ideas was truly appreciated.

To my colleagues of both groups, for their support and our time together. A special thanks to Catherine, Shahrzad, Michael, Jens, Felix, Frank, Cansu, Marco, Chris, Marin, Sebastian, Pavel, Pierre, Anselm, Max, Huanbo, Melih and Furio!

To my friends, your support and encouragement have been a big help throughout this journey.

A heartfelt thanks to Daniel Borgis and Luc Belloni for hosting me and for the productive and enjoyable time in Paris.

Finally, I'm grateful to my family, and in particular my mother and aunt. Your support and encouragement have been with me every step of the way.

I've seen things you people wouldn't believe. Attack ships on fire off the shoulder of Orion. I watched C-beams glitter in the dark near the Tannhäuser Gate. All those moments will be lost in time, like tears in rain. Time to die.
— Roy Batty, *Blade Runner (1982)*

Bibliography

[AT17] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, 2017. DOI: 10.1093/oso/9780198803195.001.0001.

[Bax68] R. J. Baxter. "Percus–Yevick Equation for Hard Spheres with Surface Adhesion". In: *The Journal of Chemical Physics* 49.6 (1968), pp. 2770–2774. DOI: 10.1063/1.1670482.

[Bel17] L. Belloni. "Exact molecular direct, cavity, and bridge functions in water system". In: *The Journal of Chemical Physics* 147.16 (2017), p. 164121. DOI: 10.1063/1.5001684.

[Bia+06] E. Bianchi et al. "Phase Diagram of Patchy Colloids: Towards Empty Liquids". In: *Physical Review Letters* 97.16 (2006), p. 168301. DOI: 10.1103/PhysRevLett.97.168301.

[Bia+08] E. Bianchi et al. "Theoretical and numerical study of the phase diagram of patchy colloids: Ordered and disordered patch arrangements". In: *The Journal of Chemical Physics* 128.14 (2008), p. 144504. DOI: 10.1063/1.2888997.

[Blu72] L. Blum. "Invariant Expansion. II. The Ornstein-Zernike Equation for Nonspherical Molecules and an Extended Solution to the Mean Spherical Model". In: *The Journal of Chemical Physics* 57.5 (1972), pp. 1862–1869. DOI: 10.1063/1.1678503.

[Blu73] L. Blum. "Invariant expansion III: The general solution of the mean spherical model for neutral spheres with electostatic interactions". In: *The Journal of Chemical Physics* 58.8 (1973), pp. 3295–3303. DOI: 10.1063/1.1679655.

[Bra+18] J. Bradbury et al. *JAX: composable transformations of Python+NumPy programs*. 2018. URL.

[BT72] L. Blum and A. J. Torruella. "Invariant Expansion for Two-Body Correlations: Thermodynamic Functions, Scattering, and the Ornstein–Zernike Equation". In: *The Journal of Chemical Physics* 56.1 (1972), pp. 303–310. DOI: 10.1063/1.1676864.

[BT88] L. Blum and A. J. Torruella. "Invariant expansion. IV. The exponentials of tensorial expressions". In: *The Journal of Chemical Physics* 89.8 (1988), pp. 4976–4980. DOI: 10.1063/1.455640.

[Cat+21] P. Cats et al. "Machine-learning free-energy functionals using density profiles from simulations". In: *APL Materials* 9.3 (2021), p. 031109. DOI: 10.1063/5.0042558.

[CGL22] M. Chang, T. Griffiths, and S. Levine. "Object Representations as Fixed Points: Training Iterative Refinement Algorithms with Implicit Differentiation". In: *Advances in Neural Information Processing Systems* 35 (2022), pp. 32694–32708. URL.

[CGS16] S. M. Cattes, K. E. Gubbins, and M. Schoen. "Mean-field density functional theory of a nanoconfined classical, three-dimensional Heisenberg fluid. I. The role of molecular anchoring". In: *The Journal of Chemical Physics* 144.19 (2016), p. 194704. DOI: 10.1063/1.4949330.

[Cho+99] C. H. Choi et al. "Rapid and stable determination of rotation matrices between spherical harmonics by direct recursion". In: *The Journal of Chemical Physics* 111.19 (1999), pp. 8825–8831. DOI: 10.1063/1.480229.

[Chu+21] M. Chupin et al. "Convergence analysis of adaptive DIIS algorithms with application to electronic ground state calculations". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 55.6 (2021), pp. 2785–2825. DOI: 10.1051/m2an/2021069.

[CJG88] W. G. Chapman, G. Jackson, and K. E. Gubbins. "Phase equilibria of associating fluids: Chain molecules with multiple bonding sites". In: *Molecular Physics* 65.5 (1988), pp. 1057–1079. DOI: 10.1080/00268978800101601.

[CTN83] E. Chacón, P. Tarazona, and G. Navascués. "A functional perturbation theory for nonuniform molecular fluids. Effect of a weak dipole in a liquid–vapor interphase". In: *The Journal of Chemical Physics* 79.9 (1983), pp. 4426–4430. DOI: 10.1063/1.446327.

[CTN84] E. Chacón, P. Tarazona, and G. Navascués. "Molecular fluids in an external field: A functional perturbation theory for the orientation profile". In: *Molecular Physics* 51.6 (1984), pp. 1475–1486. DOI: 10.1080/00268978400100971.

[Dij+24] J. Dijkman et al. *Learning Neural Free-Energy Functionals with Pair-Correlation Matching*. 2024. DOI: 10.48550/arXiv.2403.15007.

[Din+17] L. Ding et al. "Efficient molecular density functional theory using generalized spherical harmonics expansions". In: *The Journal of Chemical Physics* 147.9 (2017), p. 094107. DOI: 10.1063/1.4994281.

[Din17] L. Ding. "Molecular Density Functional Theory under homogeneous reference fluid approximation". PhD thesis. Université Paris Saclay, 2017. URL.

[Eva79] R. Evans. "The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids". In: *Advances in Physics* 28.2 (1979), pp. 143–200. DOI: 10.1080/00018737900101365.

[Fen+15] X. M. Feng et al. "High-precision evaluation of Wigner's d matrix by exact diagonalization". In: *Physical Review E* 92.4 (2015), p. 043307. DOI: 10.1103/PhysRevE.92.043307.

[Fra+24] F. O. de Franca et al. "SRBench++: Principled Benchmarking of Symbolic Regression With Domain-Expert Interpretation". In: *IEEE Transactions on Evolutionary Computation* (2024), pp. 1–1. DOI: 10.1109/TEVC.2024.3423681.

[FS07] G. Foffi and F. Sciortino. "On the Possibility of Extending the Noro–Frenkel Generalized Law of Correspondent States to Nonisotropic Patchy Interactions". In: *The Journal of Physical Chemistry B* 111.33 (2007), pp. 9702–9705. DOI: 10.1021/jp074253r.

[FS23] D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Elsevier, 2023. DOI: 10.1016/C2009-0-63921-0.

[GC93] D. Ghonasgi and W. G. Chapman. "Theory and simulation for associating fluids with four bonding sites". In: *Molecular Physics* 79.2 (1993), pp. 291–311. DOI: 10.1080/00268979300101221.

[Gen+21] Z. Geng et al. "On Training Implicit Models". In: *Advances in Neural Information Processing Systems*. Vol. 34. 2021, pp. 24247–24260. URL.

[GG84] C. G. Gray and K. E. Gubbins. *Theory of Molecular Fluids: Volume I: Fundamentals*. Oxford University Press, 1984. DOI: 10.1093/oso/9780198556022.001.0001.

[GGJ11] C. G. Gray, K. E. Gubbins, and C. G. Joslin. *Theory of Molecular Fluids: Volume 2: Applications*. Oxford University Press, 2011. DOI: 10.1093/acprof:oso/9780198556213.001.0001.

[Gna+12] N. Gnan et al. "Properties of patchy colloidal particles close to a surface: A Monte Carlo and density functional study". In: *The Journal of Chemical Physics* 137.8 (2012), p. 084704. DOI: 10.1063/1.4746428.

[Ham89] M. Hamermesh. *Group Theory and its Application to Physical Problems*. Dover Publications, 1989.

[HM13] J.-P. Hansen and I. R. McDonald. *Theory of Simple Liquids: With Applications to Soft Matter*. Academic Press, 2013. DOI: 10.1016/C2010-0-66723-X.

[HR06] H. Hansen-Goos and R. Roth. "Density functional theory for hard-sphere mixtures: the White Bear version mark II". In: *Journal of Physics: Condensed Matter* 18.37 (2006), p. 8413. DOI: 10.1088/0953-8984/18/37/002.

[HTG11] D. de las Heras, J. M. Tavares, and M. M. T. da Gama. "Phase diagrams of binary mixtures of patchy colloids with distinct numbers and types of patches: The empty fluid regime". In: *The Journal of Chemical Physics* 134.10 (2011), p. 104904. DOI: 10.1063/1.3561396.

[JCG88] G. Jackson, W. G. Chapman, and K. E. Gubbins. "Phase equilibria of associating fluids: Spherical molecules with multiple bonding sites". In: *Molecular Physics* 65.1 (1988), pp. 1–31. DOI: 10.1080/00268978800100821.

[JF16] H. T. Johansson and C. Forssén. "Fast and Accurate Evaluation of Wigner 3*j*, 6*j*, and 9*j* Symbols Using Prime Factorization and Multiword Integer Arithmetic". In: *SIAM Journal on Scientific Computing* 38.1 (2016), A376–A384. DOI: 10.1137/15M1021908.

[KF03] N. Kern and D. Frenkel. "Fluid–fluid coexistence in colloidal systems with short-ranged strongly directional attraction". In: *The Journal of Chemical Physics* 118.21 (2003), pp. 9882–9889. DOI: 10.1063/1.1569473.

[Kuf04] J. Kuffner. "Effective sampling and distance metrics for 3D rigid body path planning". In: *IEEE International Conference on Robotics and Automation*. Vol. 4. 2004, pp. 3993–3998. DOI: 10.1109/ROBOT.2004.1308895.

[LLL95] F. Lado, E. Lomba, and M. Lombardero. "Integral equation algorithm for fluids of fully anisotropic molecules". In: *The Journal of Chemical Physics* 103.1 (1995), pp. 481–484. DOI: 10.1063/1.469615.

[LMO20] S.-C. Lin, G. Martius, and M. Oettel. "Analytical classical density functionals from an equation learning network". In: *The Journal of Chemical Physics* 152.2 (2020), p. 021102. DOI: 10.1063/1.5135919.

[LO19] S.-C. Lin and M. Oettel. "A classical density functional from machine learning and a convolutional neural network". In: *SciPost Physics* 6.2 (2019), p. 025. DOI: 10.21468/SciPostPhys.6.2.025.

[Lut10] J. F. Lutsko. "Recent Developments in Classical Density Functional Theory". In: *Advances in Chemical Physics* 144 (2010), pp. 1–92. DOI: 10.1002/9780470564318.

[Lut13] J. F. Lutsko. "Direct correlation function from the consistent fundamental-measure free energies for hard-sphere mixtures". In: *Physical Review E* 87.1 (2013), p. 014103. DOI: 10.1103/PhysRevE.87.014103.

[LWK18] C. Louizos, M. Welling, and D. P. Kingma. *Learning Sparse Neural Networks through L_0 Regularization*. 2018. DOI: 10.48550/arXiv.1712.01312.

[Ma+22] H. Ma et al. "Evolving symbolic density functionals". In: *Science Advances* 8.36 (2022), eabq0279. DOI: 10.1126/sciadv.abq0279.

[Mar14] B. D. Marshall. "Thermodynamic Perturbation Theory for Associating Fluids: Beyond First Order". PhD thesis. Rice University, 2014. URL.

[MC16] B. D. Marshall and W. G. Chapman. "Thermodynamic Perturbation Theory for Associating Molecules". In: *Advances in Chemical Physics*. John Wiley & Sons, 2016, pp. 1–47. DOI: 10.1002/9781119165156.ch1.

[MCT13] B. D. Marshall, W. G. Chapman, and M. M. Telo da Gama. "Classical density functional theory for associating fluids in orienting external fields". In: *Physical Review E* 88.6 (2013), p. 060301. DOI: 10.1103/PhysRevE.88.060301.

[Mil65] R. E. Miles. "On Random Rotations in R^3". In: *Biometrika* 52.3/4 (1965), pp. 636–639. DOI: 10.2307/2333716.

[ML16] G. Martius and C. H. Lampert. *Extrapolation and learning equations*. 2016. DOI: 10.48550/arXiv.1610.02995.

[MMT17] C. J. Maddison, A. Mnih, and Y. W. Teh. *The Concrete Distribution: A Continuous Relaxation of Discrete Random Variables*. 2017. DOI: 10.48550/arXiv.1611.00712.

[RD19] P. Ramadevi and V. Dubey. *Group Theory for Physicists: With Applications*. Cambridge: Cambridge University Press, 2019. DOI: 10.1017/9781108554862.

[Ros+97] Y. Rosenfeld et al. “Fundamental-measure free-energy density functional for hard spheres: Dimensional crossover and freezing”. In: *Physical Review E* 55.4 (1997), pp. 4245–4263. DOI: 10.1103/PhysRevE.55.4245.

[Ros89] Y. Rosenfeld. “Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing”. In: *Physical Review Letters* 63.9 (1989), pp. 980–983. DOI: 10.1103/PhysRevLett.63.980.

[Rot+02] R. Roth et al. “Fundamental measure theory for hard-sphere mixtures revisited: the White Bear version”. In: *Journal of Physics: Condensed Matter* 14.46 (2002), p. 12063. DOI: 10.1088/0953-8984/14/46/313.

[Rot10] R. Roth. “Fundamental measure theory for hard-sphere mixtures: a review”. In: *Journal of Physics: Condensed Matter* 22.6 (2010), p. 063102. DOI: 10.1088/0953-8984/22/6/063102.

[RRR18] L. Rovigatti, J. Russo, and F. Romano. “How to simulate patchy particles”. In: *The European Physical Journal E* 41.5 (2018), p. 59. DOI: 10.1140/epje/i2018-11667-x.

[RSS10] F. Romano, E. Sanz, and F. Sciortino. “Phase diagram of a tetrahedral patchy particle model for different interaction ranges”. In: *The Journal of Chemical Physics* 132.18 (2010), p. 184501. DOI: 10.1063/1.3393777.

[Rus+22] J. Russo et al. “The physics of empty liquids: from patchy particles to water”. In: *Reports on Progress in Physics* 85.1 (2022), p. 016601. DOI: 10.1088/1361-6633/ac42d9.

[Sam+23] F. Sammüller et al. “Neural functional theory for inhomogeneous fluids: Fundamentals and applications”. In: *Proceedings of the National Academy of Sciences* 120.50 (2023), e2312484120. DOI: 10.1073/pnas.2312484120.

[San+14] T. Santos-Silva et al. “Neural-network approach to modeling liquid crystals in complex confinement”. In: *Physical Review E* 89.5 (2014), p. 053316. DOI: 10.1103/PhysRevE.89.053316.

[San16] A. Santos. *A Concise Course on the Theory of Classical Liquids: Basics and Selected Topics*. Vol. 923. Lecture Notes in Physics. Cham: Springer International Publishing, 2016. DOI: 10.1007/978-3-319-29668-5.

[Sci08] F. Sciortino. “Gel-forming patchy colloids and network glass formers: thermodynamic and dynamic analogies”. In: *The European Physical Journal B* 64.3 (2008), pp. 505–509. DOI: 10.1140/epjb/e2008-00034-0.

[SCS97] C. Segura, W. Chapman, and K. Shukla. "Associating fluids with four bonding sites against a hard wall: density functional theory". In: *Molecular Physics* 90.5 (1997), pp. 759–771. DOI: 10.1080/00268979709482661.

[SGK14] M. Schoen, S. Giura, and S. H. L. Klapp. "Phase behavior of an amphiphilic fluid". In: *Physical Review E* 89.1 (2014), p. 012310. DOI: 10.1103/PhysRevE.89.012310.

[Sim+24] A. Simon et al. "Machine Learning of a Density Functional for Anisotropic Patchy Particles". In: *Journal of Chemical Theory and Computation* 20.3 (2024), pp. 1062–1077. DOI: 10.1021/acs.jctc.3c01238.

[Sim+25] A. Simon et al. "The orientational structure of a model patchy particle fluid: Simulations, integral equations, density functional theory, and machine learning". In: *The Journal of Chemical Physics* 162.3 (2025), p. 034503. DOI: 10.1063/5.0248694.

[SLM18] S. S. Sahoo, C. H. Lampert, and G. Martius. *Learning Equations for Extrapolation and Control*. 2018. DOI: 10.48550/arXiv.1806.07259.

[SO24] A. Simon and M. Oettel. *Machine Learning approaches to classical density functional theory*. 2024. DOI: 10.48550/arXiv.2406.07345.

[Sol13] J. R. Solana. *Perturbation Theories for the Thermodynamic Properties of Fluids and Solids*. CRC Press, 2013. DOI: 10.1201/b13960.

[SS24] F. Sammüller and M. Schmidt. "Neural density functionals: Local learning and pair-correlation matching". In: *Physical Review E* 110.3 (2024), p. L032601. DOI: 10.1103/PhysRevE.110.L032601.

[Ste80] W. A. Steele. "Symmetry constraints on the configurational properties of non-linear molecules". In: *Molecular Physics* 39.6 (1980), pp. 1411–1422. DOI: 10.1080/00268978000101151.

[Sti+20] R. Stierle et al. "Guide to efficient solution of PC-SAFT classical Density Functional Theory in various Coordinate Systems using fast Fourier and similar Transforms". In: *Fluid Phase Equilibria* 504 (2020), p. 112306. DOI: 10.1016/j.fluid.2019.112306.

[Sti21] R. Stierle. "Development of hydrodynamic density functional theory for mixtures and application to droplet coalescence". PhD thesis. Universität Stuttgart, 2021. DOI: 10.18419/opus-11890.

[Sto+18] D. Stopper et al. "Bulk structural information from density functionals for patchy particles". In: *The Journal of Chemical Physics* 149.22 (2018), p. 224503. DOI: 10.1063/1.5064780.

[Sto+20] D. Stopper et al. “Remnants of the disappearing critical point in chain-forming patchy fluids”. In: *The Journal of Chemical Physics* 152.11 (2020), p. 111101. DOI: 10.1063/1.5141059.

[Sto19] D. Stopper. “Structure and Dynamics of Model Fluids with Anisotropic Interactions”. PhD thesis. Universität Tübingen, 2019. DOI: 10.15496/publikation-36484.

[Tar00] P. Tarazona. “Density Functional for Hard Sphere Crystals: A Fundamental Measure Approach”. In: *Physical Review Letters* 84.4 (2000), pp. 694–697. DOI: 10.1103/PhysRevLett.84.694.

[Tar85] P. Tarazona. “Free-energy density functional for hard spheres”. In: *Physical Review A* 31.4 (1985), pp. 2672–2679. DOI: 10.1103/PhysRevA.31.2672.

[Tav+95] J. M. Tavares et al. “Phase diagram and critical behavior of the ferromagnetic Heisenberg fluid from density-functional theory”. In: *Physical Review E* 52.2 (1995), pp. 1915–1929. DOI: 10.1103/PhysRevE.52.1915.

[TG91] P. I. Teixeira and M. M. T. d. Gama. “Density-functional theory for the interfacial properties of a dipolar fluid”. In: *Journal of Physics: Condensed Matter* 3.1 (1991), p. 111. DOI: 10.1088/0953-8984/3/1/009.

[TN82] P. Tarazona and G. Navascués. “A perturbation-variational theory for molecular fluid interphases”. In: *Molecular Physics* 47.1 (1982), pp. 145–160. DOI: 10.1080/00268978200100102.

[TS19] P. I. C. Teixeira and F. Sciortino. “Patchy particles at a hard wall: Orientation-dependent bonding”. In: *The Journal of Chemical Physics* 151.17 (2019), p. 174903. DOI: 10.1063/1.5124008.

[VMK88] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii. *Quantum Theory of Angular Momentum*. World Scientific, 1988. DOI: 10.1142/0270.

[Wer84a] M. S. Wertheim. “Fluids with highly directional attractive forces. I. Statistical thermodynamics”. In: *Journal of Statistical Physics* 35.1 (1984), pp. 19–34. DOI: 10.1007/BF01017362.

[Wer84b] M. S. Wertheim. “Fluids with highly directional attractive forces. II. Thermodynamic perturbation theory and integral equations”. In: *Journal of Statistical Physics* 35.1 (1984), pp. 35–47. DOI: 10.1007/BF01017363.

[Wer86a] M. S. Wertheim. "Fluids with highly directional attractive forces. III. Multiple attraction sites". In: *Journal of Statistical Physics* 42.3 (1986), pp. 459–476. DOI: 10.1007/BF01127721.

[Wer86b] M. S. Wertheim. "Fluids with highly directional attractive forces. IV. Equilibrium polymerization". In: *Journal of Statistical Physics* 42.3 (1986), pp. 477–492. DOI: 10.1007/BF01127722.

[WRS18] S. M. Wandrei, R. Roth, and M. Schoen. "Mean-field density functional theory of a nanoconfined classical, three-dimensional Heisenberg fluid. II. The interplay between molecular packing and orientational order". In: *The Journal of Chemical Physics* 149.5 (2018), p. 054704. DOI: 10.1063/1.5040934.

[Yan+91] B. Yang et al. "Molecular orientational structure of the water liquid/-vapour interface". In: *Journal of Physics: Condensed Matter* 3.42 (1991), F109. DOI: 10.1088/0953-8984/3/42/010.

[Yan+92] B. Yang et al. "Density-functional theory of the water liquid–vapour interface". In: *Molecular Physics* 76.3 (1992), pp. 709–735. DOI: 10.1080/00268979200101631.

[Yer+10] A. Yershova et al. "Generating Uniform Incremental Grids on $SO(3)$ Using the Hopf Fibration". In: *The International Journal of Robotics Research* 29.7 (2010), pp. 801–812. DOI: 10.1177/0278364909352700.

[YW02a] Y.-X. Yu and J. Wu. "A fundamental-measure theory for inhomogeneous associating fluids". In: *The Journal of Chemical Physics* 116.16 (2002), pp. 7094–7103. DOI: 10.1063/1.1463435.

[YW02b] Y.-X. Yu and J. Wu. "Structures of hard–sphere fluids from a modified fundamental-measure theory". In: *The Journal of Chemical Physics* 117.22 (2002), pp. 10156–10164. DOI: 10.1063/1.1520530.

[ZG16] W. Zmpitas and J. Gross. "Detailed pedagogical review and analysis of Wertheim's thermodynamic perturbation theory". In: *Fluid Phase Equilibria*. Theo W. de Loos Festschrift 428 (2016), pp. 121–152. DOI: 10.1016/j.fluid.2016.07.033.

[Zho+19] Y. Zhou et al. "On the Continuity of Rotation Representations in Neural Networks". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019, pp. 5745–5753. URL.