From Particle Condensation to Polymer Aggregation:

Phase Transitions and Structural Phases in Mesoscopic Systems

Von der Fakultät für Physik und Geowissenschaften
der Universität Leipzig
genehmigte

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt von M.Sc. Johannes Zierenberg
geboren am 11.04.1985 in Göttingen.

Gutachter: Prof. Dr. Friederike Schmid (Universität Mainz)
Prof. Dr. Wolfhard Janke (Universität Leipzig)

Tag der Verleihung: 25.01.2016

Contents

List of Figures

List of Tables

1. Introduction

Understanding the mechanisms of particle condensation and polymer aggregation in finite systems is of relevance for a wide range of research. Advances towards applications on the nanoscale create both the need for a deeper understanding of transition mechanisms and possibilities for experimental tests of theoretical predictions on small scales and in confinement. Moreover, the formation of polymer aggregates is a basic process in biophysical networks and an underlying mechanism for protein aggregation. In general, these systems are very complex with a large number of degrees of freedom.

A common approach in physics is the reduction of a problem to its essentials. This allows one to study complex systems and to identify generic mechanisms, even when the *tough reality* is not accessible. If the gained insights may be replaced by an effective description, the next level of complexity comes into range. Besides, trying instead to study the full complex problem with many unknowns may obscure otherwise obvious key processes. Furthermore, the consideration of essentials allows one to find and study equivalences and symmetries in the description of nature, ranging from elementary particle interactions to the equivalence between an Ising magnet and a particle gas. The latter is a prominent example of universality in classical statistical physics. Here, one is usually interested in the so-called thermodynamic limit, or in other words the limit of infinite system size. However, numerical methods usually only allow the treatment of finite systems, i.e., finite volume and number of particles. In that case, finite-size scaling [1] is a useful tool to study these problems in the context of phase transitions and critical phenomena.

In this spirit, particles and polymers may be described by coarse-grained models, simplifying complex details and thereby reducing the degrees of freedoms. This includes the simplification of structural details and the reduction of complex interactions to effective pair-potentials. In general, polymers are large molecules made up of small units called monomers. In the simplest form, they appear as linear chains of identical monomers (homopolymers) or different types of monomers (heteropolymers). However, different geometries such as branched structures or (closed) rings are also possible. A coarse-grained, abstract formulation of polymers allowed the derivation of scaling relations for single flexible homopolymers and polymer solutions [2, 3], which have been verified both numerically and in experiment. Examples of linear polymers are synthetic polymers, like polyethylene, polyvinyl chloride (PVC), and polystyrene, commonly known as plastics. These may be successfully described by standard models of flexible polymers. In a completely different context, in living organisms, biopolymers assume important functions. One example are proteins, which have a heterogeneous primary structure of amino acids. This gives rise to secondary structures such as alpha helices and beta sheets, which may provide a given functionality if folded correctly. On the other hand, the aggregation of misfolded proteins is associated with several human diseases like Alzheimer's, Parkinson's, and diabetes II [4]. In the context of protein aggregation, different motifs are possible, e.g., amorphous aggregates and twisted fibrils for amyloid aggregation. In some cases, proteins assemble into homogeneous chains, where they become monomeric units. Examples are microtubules, intermediate filaments, and actin filaments, which make up the scaffold of a cell and may be

described as rather stiff linear polymers in coarse-grained networks.

Recent advances in experimental and simulation techniques open the door to mesoscopic length scales [5–7]. This enables tests of fundamental concepts of statistical physics, e.g., by predicting and verifying finite-size corrections for microscopic models and checking them experimentally. Moreover, statistical and dynamic properties of mesoscopic systems may be studied using a microscopic description to probe the models for consistency and to yield predictions. Of common interest are structural properties, the effect of confinement, and finite-size transitions in general, particularly due to an increasing experimental interest in the nanoscale regime with regard to potential applications. One interesting aspect is the competition of generic mechanisms, which may be numerically studied from small to mesoscopic system sizes. This approach also allows the study of equivalences and relations, for example the similarities of crystalline low-temperature structures of a single flexible polymer to nanoclusters.

The aim of this research is to investigate the aggregation transition of semiflexible polymers in dilute solution, with a focus on structural phases and the analogy to condensation of a particle gas. The conclusions are based on theoretical arguments and results from advanced Monte Carlo simulation techniques. I follow a bottom-up approach, from particle condensation to semiflexible polymer aggregation, in chapters 2-4. These chapters build on each other, starting with the "simpler" case of the equilibrium condensation/evaporation transition of particles in chapter 2, which is treated in detail. This includes a review and extension of the relevant theory, supported by numerical results, as well as an introduction to finite-size scaling in the microcanonical ensemble and scaling of the free-energy barrier. Chapter 3 is concerned with the aggregation transition of flexible polymers in different scaling limits, in comparison to the finite-size scaling of particle condensation. In addition, the dependence of the aggregation transition on density is studied in spherical confinement, and the effects of unequal inter- and intra-polymer interactions are discussed. The model is extended to semiflexible polymers in chapter 4, where the role of stiffness on the resulting aggregate motifs and consequences for finite-size scaling or in spherical confinement are discussed. The numerical methods used for the generation of the supporting data are described in chapter 5. This includes the description and discussion of a novel parallel implementation of the multicanonical method. Chapter 6 summarizes the main conclusions and includes a brief outlook.

2. Condensation/evaporation transition in liquid-vapor systems

The equilibrium condensation/evaporation transition separates a gas (or vapor) phase from a mixed phase of a single macroscopic droplet in equilibrium with surrounding gas. This is a good example of a first-order phase transition either maximizing entropy in the gas phase or minimizing energy in the droplet phase. Early studies of equilibrium condensation include the analytical work by Fisher [8]. Of course, considering the kinetic process of nucleation initially leads to a distribution of droplet sizes in a metastable state that eventually reach equilibrium after a (possibly very long) lag time (see e.g. Ref. [9]). The topic of equilibrium droplets includes a long list of theoretical works among others on the leading-order scaling behavior of equilibrium droplets [10–13] with origins already in the 80s [14]. The common approach to this problem is to consider a fixed temperature, studying the transition density above which a supersaturated gas forms a droplet (see Sec 2.2). Biskup et al. [10, 11] proposed a general (idealized) theory claiming that intermediate-sized droplets have a vanishing probability, with a rigorous analytical solution for the two-dimensional Ising spin model. Numerous simulation studies have addressed the condensation/evaporation at fixed temperature, including studies of the two- and three-dimensional Ising (lattice gas) model [12, 15–17] and the three-dimensional Lennard-Jones gas model [18, 19]. Equivalently, one may approach the condensation/evaporation transition by considering a fixed density, identifying a transition temperature. In Sec 2.3, the corresponding finite-size scaling of the transition temperature and rounding of the transition will be discussed. This has been demonstrated partially for the two- and three-dimensional lattice gas [20] and is, in this thesis, reconfirmed including also the Lennard-Jones gas [21]. In addition, I will present an effective intermediate scaling regime which may be successfully described by including higher-order correction terms.

This chapter is organized in the following way. In Sec. 2.1, I will introduce the employed particle gas models, discuss equivalences, and compare to exact results where possible. The main theory and results will be presented in Sec. 2.2 and Sec. 2.3. The remaining part of this chapter will be devoted to the discussion of the accompanying free-energy barrier (Sec. 2.4) and a microcanonical analysis (Sec. 2.5) in the scheme of fixed density. Also, microcanonical arguments in an ideal gas approximation are provided in order to describe the density dependence of the infinite- and finite-system transition temperature (Sec. 2.5). This will serve as a basis for the study of polymer aggregation on a more intuitive problem and will help to identify similarities between polymer aggregation and particle condensation.

2.1. Discrete and continuous particle gas models

I considered two different, simplified particle models in order to draw general conclusions, namely the discrete lattice gas (DLG) model and the Lennard-Jones gas (LJG) model. For a sketch see Fig. 2.1. This allows to distinguish between model dependent artifacts and general statements on universal aspects. The lattice gas model allows to bridge the gap to analytical

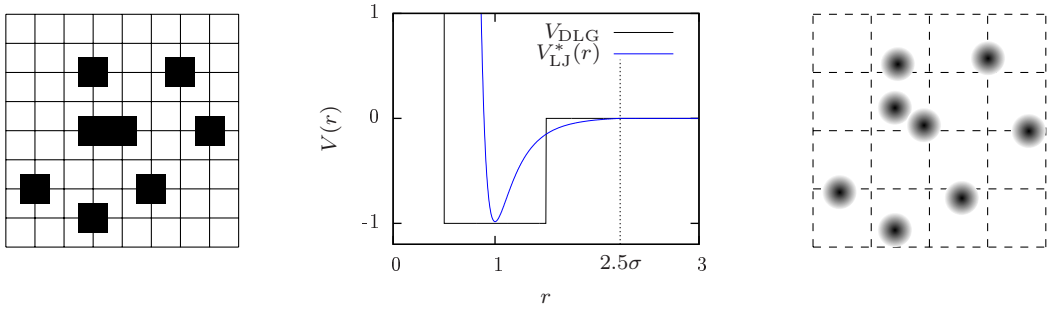

Figure 2.1.: Sketch of two complementary particle gas models: the lattice gas model (left) and the Lennard-Jones gas model with domain decomposition (right). The corresponding particle-particle interactions are demonstrated in the center.

results [10, 13], being equivalent to the Ising model at fixed magnetization. This will be exploited in two and three dimensions. In contrary to the lattice gas model, the Lennard-Jones gas is not symmetric with respect to particle-hole exchange and will be considered in three dimensions. The boundary conditions are periodic, such that interactions with boundaries are avoided. Both models have short interaction ranges by construction.

2.1.1. Lattice gas

The discrete lattice gas model (DLG) describes particles as occupied sites on a square (2D), cubic (3D) or higher-dimensional lattice. The hard-core repulsion of particles is included by the condition that each site may be either occupied or empty: $n_i \in \{0, 1\}$. The short-range interaction is modeled by nearest-neighbor interaction $\langle i, j \rangle$, yielding the Hamiltonian

$$\mathcal{H}_{\mathrm{DLG}} = -J \sum_{\langle i,j \rangle} n_i n_j. \tag{2.1}$$

The coupling constant J may be trivially included in a rescaled temperature and is thus set to $J = 1$ for all simulations. A sketch of the model is shown in Fig. 2.1 with the corresponding particle-particle interaction potential $V_{\mathrm{DLG}}(r)$, where r is the distance between two particles, including self-avoidance at $r = 0$ ($V_{\mathrm{DLG}}(0) = \infty$), nearest-neighbor attraction at $r = 1$ ($V_{\mathrm{DLG}}(1) = -1$), and no interaction else.

The natural (canonical) ensemble in this formulation is a fixed particle number N, system volume V and temperature T: the NVT ensemble. In general, the inverse thermal energy or otherwise known as inverse temperature $\beta = 1/k_{\mathrm{B}}T$ sets the energy scale, see also Sec. 5.1. However, I usually consider the temperature T in the common dimensionless units where $k_{\mathrm{B}} = 1$. Since both total number of particles $N = \sum_i n_i$ and total number of sites $V = L^d$ are integers, the density $\rho = N/V$ may only be adjusted approximately. The lattice gas model may be treated with exact enumerations for very small system sizes and of course with Markov chain Monte Carlo methods like the Metropolis and multicanonical method (for details see Ch. 5). The Monte Carlo updates I considered include single particle shifts to a random nearest neighbor and single particle jumps to a random new site, see also Sec. 5.7.1. In addition to the energy E and its thermal derivative the specific heat $\langle C_V \rangle = \frac{d}{dT}\langle E \rangle / N$, I consider another

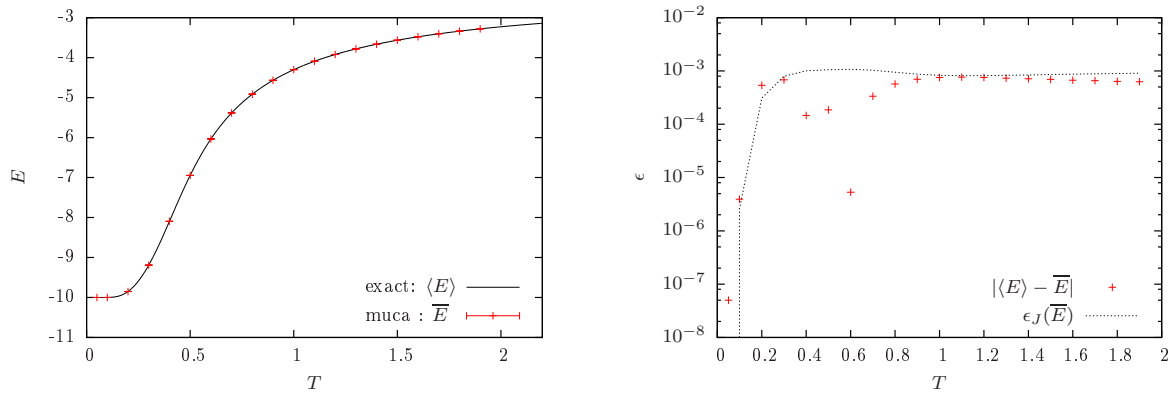

Figure 2.2.: Comparison of parallel multicanonical simulations of a two-dimensional lattice gas with $N = 8$ particles on a 7×7 square lattice with exact enumerations: the canonical total energy shows good qualitative agreement (left). A detailed comparison of the jackknife error with the deviation of the estimated energy $\overline{E}$ from the exact expectation value $\langle E \rangle$ shows that they are on the same order of magnitude (right).

observable: the number of particles in the largest cluster N_{D}. In general, this is measured as the largest connected cluster of particles using a sequential cluster identification, merging connected clusters. This directly measures the number of particles in the cluster (or the mass). When comparing with literature, this may be extended by a flood-fill routine that measures the volume of the cluster (up to surface corrections), for details see Ref. [16].

Exact enumerations allow to validate the implementation of the model, the simulation technique, and the analysis. The exact enumeration yields the canonical expectation value $\langle O \rangle$, while an importance-sampling Markov chain only allows to calculate estimators $\overline{O}$, see also Sec. 5.1. The final data of a parallel multicanonical simulation may be reweighted to any temperature whose canonical histogram is covered by the flat histogram, yielding estimators of canonical expectation values. Then, the jackknife error ϵ_J may be estimated as described in Sec. 5.8. For the total energy, Fig. 2.2 shows the comparison of the expectation value from exact enumerations and the estimator with jackknife errors from parallel multicanonical simulations using $p = 20$ cores and 2.56 million measurements. The left panel nicely shows the qualitative agreement. In addition, the right panel shows that the estimated errors are of the same order as the absolute deviations from the exact data. Since the reweighted estimators are obtained from the same set of data, the individual estimators are correlated which can be also seen in the deviation from the exact data. The test has been performed with the identical implementation used in Sec. 2.3 and beyond. The same holds for the implementation of the analysis, which was used to obtain most data in this thesis in a generalized version.

Figure 2.3 shows the specific heat normalized with the number of particles from exact enumerations together with the corresponding estimators from the above multicanonical simulations. Again the qualitative and quantitative agreement validate the (independent) implementations of model, method, and analysis. The absolute errors of both the specific heat and the energy seem to be of the same order, while comparing the scales of the expectation values show that, in fact, the relative error of the specific heat in the vicinity of the peak ($T = 0.4$) is larger by a factor of 5. This is a good illustration of the increase in uncertainty with every

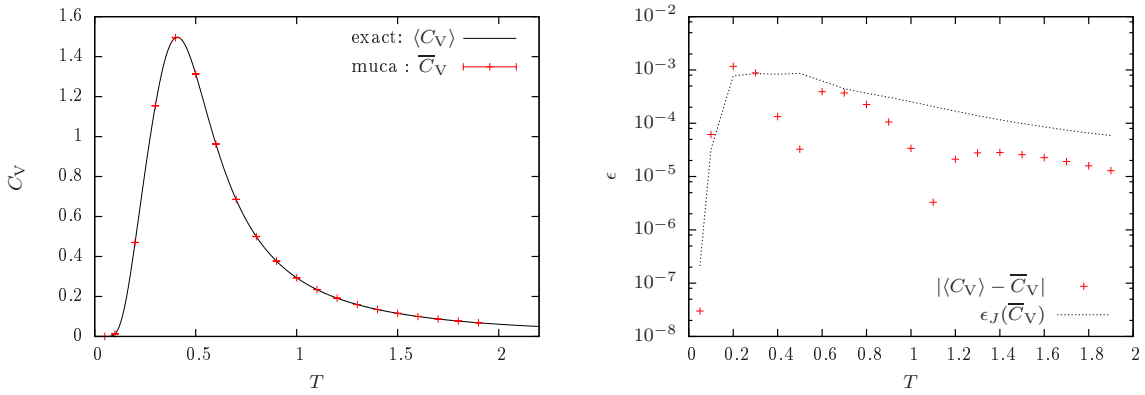

Figure 2.3.: Same as Fig. 2.2 for the specific heat estimator $\overline{C}_V$ and exact expectation value $\langle C_V \rangle$.

derivative.

Equivalence to the Ising model: This formulation of a discretized particle gas is equivalent to the Ising model (with spins $s_i \in \{\pm 1\}$) at fixed magnetization $M = \sum_i s_i$ [22]. Here, the canonical ensemble is to consider a fixed number of spins V (volume) and temperature where the magnetization becomes an observable. Thus, the canonical Ising model becomes a grand-canonical approach in the particle picture. Fixing the magnetization on the other hand leads to the desired equivalence, while loosing a natural meaning in the Ising picture. Identifying $s_i = 2n_i - 1$ and keeping $\sum_i n_i = N$ constant, one may rewrite the Ising Hamiltonian in terms of Eq. (2.1)

$$\mathcal{H}^{\mathrm{Is}} = -J \sum_{\langle i,j \rangle} s_i s_j = 4\mathcal{H}^{\mathrm{DLG}} - 2J(V - 4N). \tag{2.2}$$

The temperature scales may be mapped onto each other by considering a rescaled temperature $T' = T^{\mathrm{Is}}/4$. Actually, this factor becomes relevant for all energy-related observables and constants (for example the surface free energy, see below). The magnetization per spin $m = M/V$ in the Ising model is then related to the density of the particle gas via

$$m = 1 - 2\rho. \tag{2.3}$$

Figure 2.4 shows the equivalence of the Ising gas and lattice gas model on the example of the equilibrium number of particles in the largest droplet (cluster) N_{D}.

In two dimension, this equivalence allows to draw on some quantities which are known analytically and others which are known to arbitrary precision. This will become relevant when comparing finite-size extrapolations or predictions from leading-order results. For the 2D Ising model, a temperature-driven second-order phase transition occurs at the inverse temperature

$$\beta_c^{\mathrm{Is}} = \frac{1}{2} \ln\left(1 + \sqrt{2}\right). \tag{2.4}$$

The spontaneous magnetization m_0 is described by the Onsager-Yang equation [23, 24]:

$$m_0(\beta^{\mathrm{Is}}) = \left(1 - \sinh^{-4}(2\beta^{\mathrm{Is}})\right)^{1/8} \tag{2.5}$$

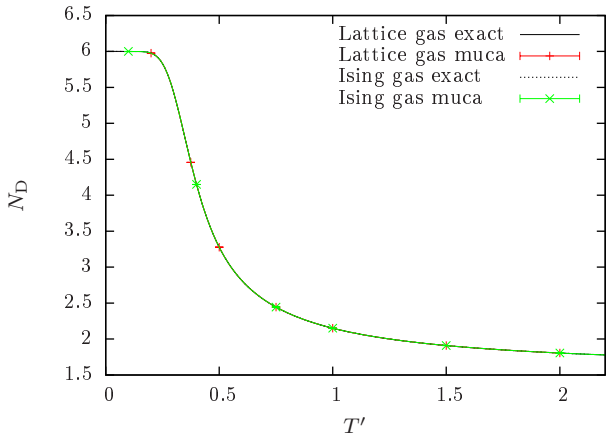

Figure 2.4.: Comparison of exact enumerations for both the two-dimensional lattice gas and Ising gas (fixed magnetization) model with corresponding multicanonical data. The system size is $L = 10$ with $N = 6$ particles. For this example, the number of particles N_D in the largest droplet (cluster) is chosen. The x-axis shows the rescaled temperature T', demonstrating the equivalence of the lattice gas and the Ising gas model for $T' = T$ and $T' = T^{Is}/4$, respectively.

The magnetic susceptibility χ may be evaluated from sufficiently long series expansions [25–30], where

$$\chi(\beta^{Is}) = \beta \sum_{i=0}^{n} c_i u^{2i} \quad \text{with} \quad u = \frac{1}{2\sinh(2\beta^{Is})}, \tag{2.6}$$

and $c = \{0, 0, 4, 16, 104, 416, 2224, 8896, 43840, 175296, 825648, 3300480, 15101920, ...\}$[1], which I used up to the 300^{th} term.
The volume of the Wulff plot or the Wulff shape (the equilibrium shape of a 2D Ising droplet) is given by [31]

$$W = \frac{4}{(\beta^{Is})^2} \int_0^{\beta^{Is}\sigma_0} dx \ \cosh^{-1}\left[\frac{\cosh^2(2\beta^{Is})}{\sinh(2\beta^{Is})} - \cosh(\beta^{Is})\right], \tag{2.7}$$

where $\sigma_0 = 2 + \ln[\tanh(\beta^{Is})]/\beta^{Is}$ and $\cosh^{-1}$ is referring to the inverse hyperbolic cosine sometimes also denoted Arccosh. This will be relevant for the surface free energy of a (Wulff shaped) droplet of unit volume $\tau_W^{Is} = 2\sqrt{W}$. Being energy-related, the interface tension gets converted as $\tau_W = \tau_W^{Is}/4$

In three dimensions, there exist suitable low-temperature series expansions which provide estimates around and below the roughening transition $T_R^{Is} \approx 2.4537$ [32, 33]. For example, the spontaneous magnetization (for zero field) may be computed from the series expansion [34, 35]

$$m_0(\beta^{Is}) = 1 - 2\left(\sum_n m_n \ \left(e^{-2\beta^{Is}}\right)^n\right), \tag{2.8}$$

with coefficients available up to order $n = 42$.

[1] The coefficients were obtained from `http://www.ms.unimelb.edu.au/~iwan/ising/Ising_ser.html` [29].

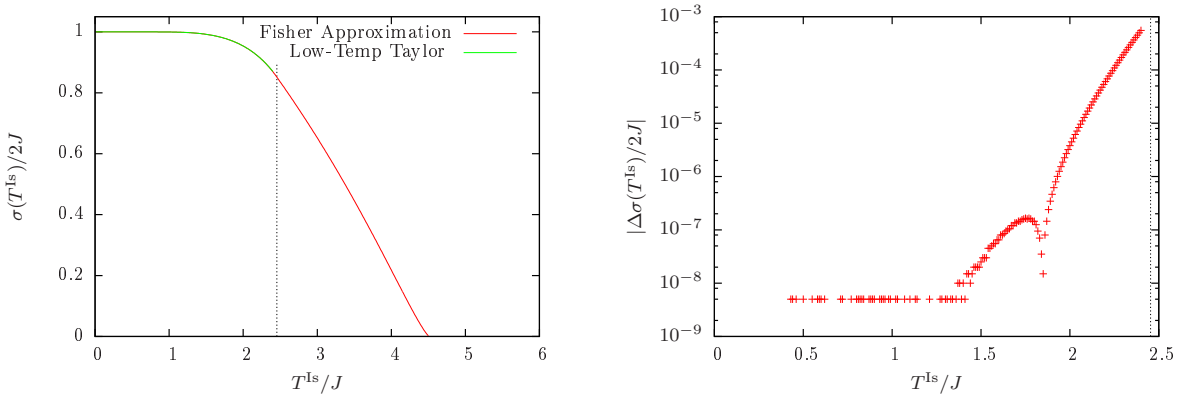

Figure 2.5.: (left) Low-temperature series expansion (2.9) and closed-form (2.10) of the interface tension $\sigma(T^{\mathrm{Is}})$. (right) Absolute difference between the two expansions. The vertical dashed line denotes the roughening temperature $T_{\mathrm{R}}^{\mathrm{Is}}$.

For the interface (surface) tension σ, there exists a low-temperature series expansions [36, 37] ($T \leq T_{\mathrm{R}}^{\mathrm{Is}}$) with a closed-form approximation. Introducing $u = \sqrt{y} = \sqrt{\exp(-4J/k_{\mathrm{B}}T^{\mathrm{Is}})}$, the low-temperature expansion may be written

$$\sigma(T^{\mathrm{Is}})/2J = 1 - \sum_n a_n \frac{u^n}{\ln(u)}, \tag{2.9}$$

with $a_4 = -2$,$a_6 = -2$,$a_8 = -10$, ... listed in Ref. [37]. On the other hand, Shaw and Fisher introduced a closed-form [36]

$$\sigma(T^{\mathrm{Is}}) \simeq \begin{cases} 2J + \frac{k_{\mathrm{B}}T^{\mathrm{Is}}}{2} \ln\left(\frac{1-0.387512y-12.2551y^2+1.96894y^3+23.4669y^4}{1-0.387512y-8.2551y^2+4.41889y^3+0.896454y^4-3.99498y^5}\right) & \text{if } y \leq 0.195 \\ 8.1301Jt^{1.264}/(1+1.6872t) & \text{else,} \end{cases} \tag{2.10}$$

where $t = (T_c^{\mathrm{Is}} - T^{\mathrm{Is}})/T_c^{\mathrm{Is}}$, and $k_{\mathrm{B}}T_c^{\mathrm{Is}}/J \approx 4.5121$ [38] the second-order temperature-driven phase transition temperature in three dimensions. For a comparison of both expansions see Fig. 2.5. Up to $k_{\mathrm{B}}T^{\mathrm{Is}}/J \approx 2$ the absolute difference between both expansions is below 10^{-6}, which provides a reasonable justification to apply the expansion in this temperature range.

2.1.2. Lennard-Jones gas

A continuous particle gas model is the Lennard-Jones gas (LJG), where in principle all particles interact with each other via the potential

$$V_{\mathrm{LJ}}(r_{ij}) = 4\epsilon\left(\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right) = \epsilon\left(\left(\frac{r_0}{r_{ij}}\right)^{12} - 2\left(\frac{r_0}{r_{ij}}\right)^{6}\right), \tag{2.11}$$

where r_{ij} is the distance between particle i and j, and $\sigma = 2^{-1/6}r_0$. In order to be comparable to the lattice gas model, I choose $\epsilon = 1$ and the potential minimum at $r_0 = 1$, which corresponds to $\sigma = 2^{-1/6}$. The computational demand may be reduced by introducing a cutoff radius $r_c = 2.5\sigma$ above which particles do not interact anymore. The potential is then shifted by

$V_{\mathrm{LJ}}(r_c)$ in order to be continuous, yielding

$$V^*_{\mathrm{LJ}}(r) = \begin{cases} V_{\mathrm{LJ}}(r) - V_{\mathrm{LJ}}(r_c) & r < r_c \\ 0 & \text{else} \end{cases}. \tag{2.12}$$

This is consistent with the existing literature and allows to apply a domain decomposition, where the periodic box is decomposed into equally large (cubic) domains. These domains have to be at least of the size r_c. Then, the interaction of each particle may be obtained by evaluating only its domain and the adjacent ones (in three dimensions this adds up to $3^3 = 27$ domains). Especially in the gas phase the simulation benefits from this procedure, where the particles are equally distributed in the full box. The Hamiltonian may then be written as

$$\mathcal{H}_{\mathrm{LJG}} = \frac{1}{2} \sum_{i \neq j} V^*_{\mathrm{LJ}}(r_{ij}). \tag{2.13}$$

A sketch is shown in Fig. 2.1 together with the interaction potential.

As Monte Carlo update, I only consider the single-particle displacement with fixed update range (see Sec. 5.7.1). As for the lattice gas, again I measure the energy E, the specific heat C_V and the number of particles in the largest droplet N_{D}. For the latter observable, particles are defined to be connected if their distance is smaller than 2σ. The clusters are then identified by a loop over all particles, associating a cluster id to the particle and all connected neighbors. If connected neighbors have the same id, the clusters are merged.

In fact, the lattice gas interaction may be interpreted as an oversimplification of the Lennard-Jones potential including the main features: mutual avoidance and short-range attraction. However, the microscopic details are different enough in order to be considered complementary.

2.2. The "native" approach: crossing the phase boundary at fixed temperature

In the "native" approach to particle condensation/evaporation, the system temperature is fixed considering a supersaturated gas with variable particle number. This requires a constant temperature smaller than a critical temperature. In a grand-canonical ensemble (with fixed chemical potential μ), this would correspond to an equilibrium particle number or *background contribution* N_0. Supersaturated then refers to a canonical ensemble (fixed NVT) with even more particles $N > N_0$, i.e., with a particle excess $\delta N = N - N_0$. Below the transition density, the excess of particles goes into the gas phase, while for larger densities droplet formation occurs. In equilibrium droplet formation, the probability for intermediate-sized droplets was shown to vanish [10] and the scenario reduces to a homogeneous gas phase and an inhomogeneous phase of a droplet in equilibrium with surrounding vapor. A fixed temperature allows to describe this effect based on fixed thermal fluctuations and relate to infinite-size temperature-dependent quantities, like the isothermal compressibility and the surface free energy. In this section, I will briefly recapture the arguments by Biskup et al. [10] yielding the leading-order finite-size correction to the transition density at fixed temperature and relate to the results from Binder [13]. I will compare this with numerical results, which have been partially published in Ref. [17].

When the probability of intermediate-size droplets vanishes, the problem of condensation

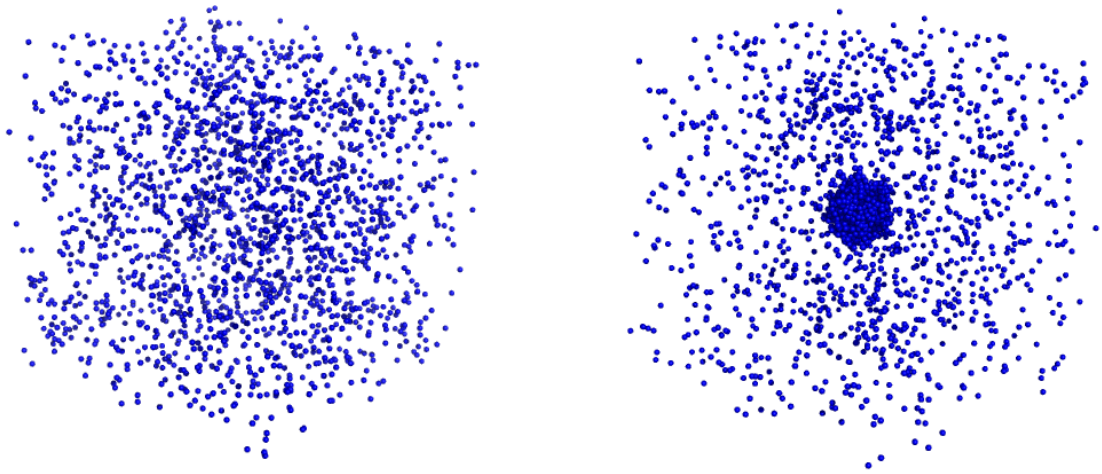

Figure 2.6.: Snapshots from Metropolis simulations of a three-dimensional Lennard-Jones gas of 2 000 particles showing (left) fluctuations of particle excess and (right) a single condensate surrounded by vapor.

may be reduced to the interplay of entropy maximization by fluctuations in the homogeneous gas phase and energy minimization by forming a single macroscopic droplet in equilibrium with surrounding bulk gas [10, 11], see also Fig. 2.6. For a supersaturated particle gas, the free energy may then be approximated by a contribution from the fluctuation of particle excess δN

$$F_{\text{fluc}} = \frac{(\delta N)^2}{2\hat{\kappa}V}, \tag{2.14}$$

with the isothermal compressibility $\hat{\kappa} = \beta\kappa = \beta \left\langle (N - \langle N \rangle)^2 \right\rangle / V$ and a contribution from the single macroscopic droplet of size V_{D}

$$F_{\text{drop}} = \tau_{\text{W}} (V_{\text{D}})^{\frac{d-1}{d}}, \tag{2.15}$$

where τ_{W} is the surface free energy of a (Wulff shaped) droplet of unit volume. These contributions are idealized with possible sources of corrections in both the Gaussian approximation and the (non-ideal) droplet shape for finite systems.

The *particle excess* was introduced as the difference between the actual particle number N and the background contribution N_0, which is the particle number expectation value in a grand-canonical ensemble at fixed temperature T and chemical potential μ (the explicit value of μ is not necessary in the following arguments). For the infinite lattice gas, the *background density* $\rho_0 = N_0/V$ of the gas phase may be identified with the canonical spontaneous magnetization $m_0(T)$ of the equivalent Ising model via Eq. (2.3). According to Biskup et al. [10], the particle excess may be decomposed into the particle excess inside the droplet δN_{D} and the particle excess in the fluctuating phase δN_{F}, i.e.,

$$\delta N = N - N_0 = \delta N_{\text{D}} + \delta N_{\text{F}}. \tag{2.16}$$

Linking the droplet size to the particle excess inside the droplet, one expects the excess $\delta N_{\mathrm{D}} = (\rho_{\mathrm{L}} - \rho_0)V_{\mathrm{D}}$, where ρ_{L} and ρ_0 are the background liquid and gas density, respectively. For the lattice gas it holds $\rho_0 = 1-\rho_{\mathrm{L}}$ due to symmetry under particle-hole exchange. Now, the two particle excesses may be further related by introducing a scalar fraction of *particle excess inside the droplet* $\lambda = \delta N_{\mathrm{D}}/\delta N$, such that $\delta N_{\mathrm{D}} = \lambda\delta N$ and $\delta N_{\mathrm{F}} = (1-\lambda)\delta N$. In simple words, λ may be understood as a normalized droplet size. The total free energy $F = F_{\mathrm{drop}} + F_{\mathrm{fluc}}$ then becomes

$$F = \left(\tau_{\mathrm{W}} \left(\frac{\lambda\delta N}{\rho_{\mathrm{L}} - \rho_0} \right)^{\frac{d-1}{d}} + \frac{(1-\lambda)^2(\delta N)^2}{2\hat{\kappa}V} \right). \tag{2.17}$$

This may be rewritten as

$$F = \tau_W \left(\frac{\delta N}{\rho_{\mathrm{L}} - \rho_0} \right)^{\frac{d-1}{d}} \left(\lambda^{\frac{d-1}{d}} + \Delta(1-\lambda)^2 \right), \tag{2.18}$$

with a dimensionless "density" parameter

$$\Delta = \frac{(\rho_{\mathrm{L}} - \rho_0)^{\frac{d-1}{d}}}{2\hat{\kappa}\tau_W} \frac{(\delta N)^{\frac{d+1}{d}}}{V}. \tag{2.19}$$

At fixed temperature $\rho_{\mathrm{L}}, \rho_0, \hat{\kappa}, \tau_W$ are constants and Δ may be interpreted as an unusual density. For a lattice gas equivalent to the Ising model, $\chi = \hat{\kappa}$ and $\tau_{\mathrm{W}}^{\mathrm{Is}} = 4\tau_W$.

This (idealized) formulation allows to identify the fraction of particles inside the (single) largest droplet λ as a function of Δ in the limit of large systems, by minimizing Eq. (2.18) with respect to λ. In fact, considering only the λ-depending part reduces the consideration to

$$\Phi_\Delta(\lambda) = \lambda^{\frac{d-1}{d}} + \Delta(1-\lambda)^2. \tag{2.20}$$

This may be minimized both analytically or numerically, yielding $\lambda(\Delta)$. It turns out (for details see Ref. [10, 11]) that there exists a constant

$$\Delta_{\mathrm{c}} = \frac{1}{d} \left(\frac{d+1}{2} \right)^{\frac{d+1}{d}}, \tag{2.21}$$

below which no condensate forms ($\lambda = 0$) and above which a single macroscopic condensate exists with non-trivial $\lambda > \lambda_{\mathrm{c}}$, see also Fig. 2.7. Directly at Δ_{c} one gets

$$\lambda_{\mathrm{c}} = \frac{2}{d+1}. \tag{2.22}$$

For the relevant dimensions considered in this thesis, this leads to $\Delta_c^{2D} \approx 0.9186$ or $\Delta_c^{3D} \approx 0.8399$ and $\lambda_c^{2D} = 2/3$ or $\lambda_c^{3D} = 1/2$, respectively.

The result $\lambda(\Delta)$ describes the expectation value of the equilibrium droplet size in the limit of large systems without any free parameter. As mentioned before, this includes the leading-order finite-size corrections with idealized assumptions. In principle, all constants may be known – in the case of the two dimensional Ising model even exactly or with very high precision. Moreover, the condensation/evaporation transition at $\Delta = \Delta_{\mathrm{c}}$ already includes the leading-order finite-size corrections to the transition density. In fact, Eq. (2.19) may be written directly in terms

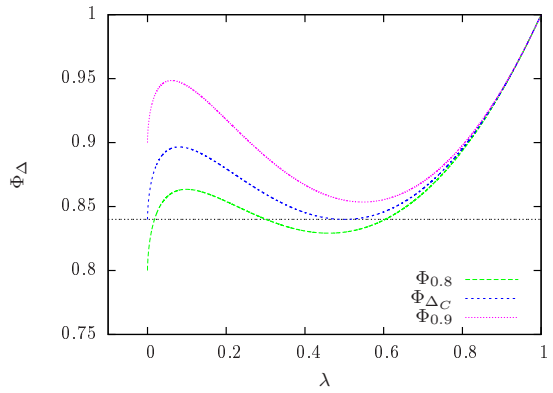

Figure 2.7.: The reduced free-energy function $\Phi_\Delta(\lambda)$ in three dimensions for several values of Δ around the transition "density" Δ_c. The minimum is either at $\lambda = 0$, namely $\Phi_\Delta(0) = \Delta$ for $\Delta < \Delta_c$ or at a non-trivial value of $\lambda > \lambda_c$.

of the (factual) particle density $\rho = N/V$

$$\Delta = \frac{(\rho_L - \rho_0)^{\frac{d-1}{d}}}{2\hat{\kappa}\tau_W} (\rho - \rho_0)^{\frac{d+1}{d}} \; V^{\frac{1}{d}}. \tag{2.23}$$

Consequently, at each finite-size transition density ρ_c it holds $\Delta(\rho_c) = \Delta_c$, which yields to leading order

$$\rho_c = \rho_0 + \left(\frac{2\hat{\kappa}\tau_W \Delta_c}{(\rho_L - \rho_0)^{\frac{d-1}{d}}} \right)^{\frac{d}{d+1}} \; V^{-\frac{1}{d+1}}. \tag{2.24}$$

In terms of the Ising model (see Sec. 2.1.1 for equivalence), this translates to a condensation magnetization

$$m_c = m_0 + 2m_0 \left(\frac{\chi \tau_W^{Is} \Delta_c}{2m_0^2} \right)^{\frac{d}{d+1}} \; V^{-\frac{1}{d+1}}. \tag{2.25}$$

This result in the notation of Biskup et al. [10, 11] is in quantitative agreement with the (independent) result of Neuhaus and Hager [12]. For the 2D Ising model, they find the same leading scaling behavior $\Delta m(L) = A_{cond} L^{-2/3}$, with the amplitude $A_{cond} = 0.23697...$ for $\beta = 0.7$. Using the constants from Sec. 2.1.1 for the same temperature yields $A = 0.236965...$ in Eq. (2.25). It is also in qualitative agreement with the results of Binder and coworkers [13, 18].

The latter works also provide an explicit leading-order scaling correction to the rounding of the transition (in this case the width $\Delta\rho$ of the finite-system transition region around ρ_c). To this end, a two-state approximation is considered where the system may be either in the condensed or evaporated phase with probability $P_{cond} \propto e^{-\beta F_{cond}}$ and $P_{evap} \propto e^{-\beta F_{evap}}$, respectively. Then, the probability to change from the homogenous state to the inhomogeneous state is related to the free-energy difference $\Delta F = F_{cond} - F_{evap}$ and the expectation value of an observable may be expressed as

$$\langle O \rangle = \frac{O_{evap} + O_{cond} \; e^{-\beta \Delta F}}{1 + e^{-\beta \Delta F}}. \tag{2.26}$$

The rounding of a transition may be related to the width defined by the condition that $|\beta\Delta F|$ is of order unity [13], which means that both phases significantly contribute to the expectation value. This is the range in which the weights of both phases differ at most by a factor e^{-1} in the expectation value. Consider the free-energy expression in Eq. (2.17). Identifying $\alpha_1 = \beta\tau_\mathrm{W}/(\rho_\mathrm{L}-\rho_0)^{\frac{d-1}{d}}$ and $\alpha_2 = \beta/2\hat{\kappa}$, the free-energy difference between the mixed phase ($\lambda \neq 0$) and the gas phase ($\lambda = 0$) may be written in terms of the density

$$\begin{aligned}\beta\Delta F(\rho)|_T &= \alpha_1\lambda^{\frac{d-1}{d}}(\rho-\rho_0)^{\frac{d-1}{d}}V^{\frac{d-1}{d}} + \alpha_2((1-\lambda)^2-1)(\rho-\rho_0)^2V \\ &= \alpha_1(\rho-\rho_0)^{\frac{d-1}{d}}V^{\frac{d-1}{d}}\underbrace{\left[\lambda^{\frac{d-1}{d}} + \frac{\alpha_2}{\alpha_1}(\rho-\rho_0)^{\frac{d+1}{d}}V^{\frac{1}{d}}\left((1-\lambda)^2-1\right)\right]}_{\Delta\Phi}. \end{aligned} \tag{2.27}$$

Expanding this around the finite-size transition density ρ_c, keeping in mind that at the transition $\beta\Delta F(\rho_\mathrm{c}) = 0$ and consequently $\Delta\Phi(\rho_\mathrm{c}) = 0$ leads in leading order to

$$\beta\Delta F(\rho)|_T = \alpha_2\left((1-\lambda)^2-1\right)\frac{d+1}{d}(\rho_\mathrm{c}-\rho_0)V(\rho-\rho_\mathrm{c}) \tag{2.28}$$

$$\propto V^{\frac{d}{d+1}}\Delta\rho, \tag{2.29}$$

where in the last line I inserted the result for ρ_c from Eq. (2.24). This was analogously derived in Refs. [13, 18] and leads for a free-energy difference of order unity to the finite-size rounding

$$\Delta\rho \propto V^{-\frac{d}{d+1}}. \tag{2.30}$$

With increasing system size, the transition width thus becomes smaller. For first-order phase transitions between homogeneous phases, this usually goes as the inverse volume V^{-1} [39].

Numerical results in two dimensions: In two dimensions, the leading-order analytical results were numerically verified in Ref. [15, 16] showing that the micromagnetic averages yield a vanishing droplet for $\Delta < \Delta_\mathrm{c}$ and a single macroscopic droplet for $\Delta > \Delta_\mathrm{c}$. Of course, finite-size effects are present and thus the functional dependence deviates slightly for finite systems. Considering for the droplet volume the largest connected cluster plus the enclosed holes, i.e., using a flood-fill routine, reduced the deviations slightly. I verified these results at $T = 1.5/4 = 0.375$ fixing the total number of particles and varying the lattice size, also considering the flood-fill definition of the droplet size $N_\mathrm{D}^\mathrm{flood}$, see Fig. 2.8. The largest cluster was measured by using a sequential cluster identification, merging connected clusters. For the rescaled plot, the following conversions are considered

$$\begin{aligned} v_\mathrm{L} &= \frac{N-\rho_0L^2}{1-2\rho_0} \\ \lambda &= N_\mathrm{D}^\mathrm{flood}/v_\mathrm{L} \\ \Delta &= \frac{(1-2\rho_0)^2}{2\chi\tau_\mathrm{W}}\frac{v_\mathrm{L}^{3/2}}{L^2} \end{aligned} \tag{2.31}$$

The overall picture is, in any case, qualitatively satisfying. The displayed analytical solution is the algebraic form from Ref. [16], using the (exact) parameters listed in Sec. 2.1.1 and is to

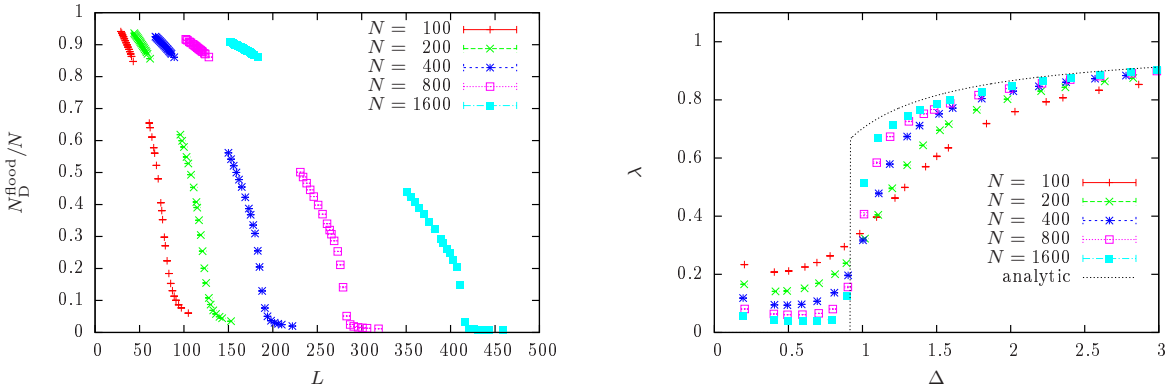

Figure 2.8.: Condensation of a lattice gas in two dimensions at $T = 0.375$: raw (left) and rescaled (right) data using multicanonical simulations.

leading-order exact. Each data point corresponds to a parallel multicanonical simulation with up to 24 cores, reweighted to the desired temperature. For lattice gas condensation, the parallel multicanonical method scales perfectly with the number of cores, as presented in Sec. 5.5.2. That way, a successful simulation was ensured not to get stuck in a local minimum but to yield a canonical estimate. At low density (or large system size) the fraction of excess in the largest droplet is small, vanishing with increasing system size, resembling the gas phase. Around the predicted transition point Δ_{c}, see Fig. 2.8 (right), the fraction of excess in the largest droplet increases and one may talk about a macroscopic droplet. With increasing system size, the functional form of the numerical data seems to approach the analytical or numerical minimization of Eq. 2.20, or $\lambda(\Delta)$. For the scope of this thesis, the two-dimensional case serves as a proof of principle. A detailed discussion of the evaporation/condensation transition and the universality with respect to the underlying lattice structure may be found in Ref. [15, 16].

Numerical results in three dimensions: In three dimensions, the comfortable situation of having exact or very precise parameters changes. In order to compare numerical data of a (Ising) lattice gas model to the leading-order analytical solution, we need m_0,χ and $\tau_{\mathrm{W}}^{\mathrm{Is}}$, see Table 2.1 for two considered temperatures. The first two parameters may be obtained from Metropolis simulations of the three dimensional Ising model at fixed temperature, measuring the average magnetization $\overline{m} = \overline{M}/V$ and the fluctuations $\overline{\chi} = \beta\left(\overline{M^2} - \overline{M}^2\right)/V$. Here, finite size corrections are small and m_0 and χ may be estimated from a constant fit to the estimates for several system sizes. Figure 2.9 shows this procedure for the example of $T^{\mathrm{Is}} = 3.2$, which corresponds to the lattice gas temperature $T = 0.8$. For the spontaneous magnetization, the low-temperature series expansion Eq. (2.8) yields $m_0^{\mathrm{low}}(T^{\mathrm{Is}} = 2.0) \approx 0.994521$ which is in good agreement with the Metropolis data. For the large temperature the expansions naturally fail with $m_0^{\mathrm{low}}(T^{\mathrm{Is}} = 3.2) \approx 0.919 - 0.927$ depending on the expansion used.

In the case of the surface free energy τ_{W}, this becomes even worse. At low temperatures, below the roughening temperature $T < T_{\mathrm{R}}$, one may use the low-temperature series expansion Eq. (2.9). For $T = 2.0$ this yields an interface tension $\sigma \approx 1.9072$. Assuming a cubic droplet shape below the roughening transition yields $\tau_{\mathrm{W}}^{\mathrm{Is}} \approx 6\sigma$. This still has to be converted to the lattice gas scale $\tau_{\mathrm{W}} = \tau_{\mathrm{W}}^{\mathrm{Is}}/4$. Moreover, τ_{W} can be obtained analogous to Ref. [16] directly from

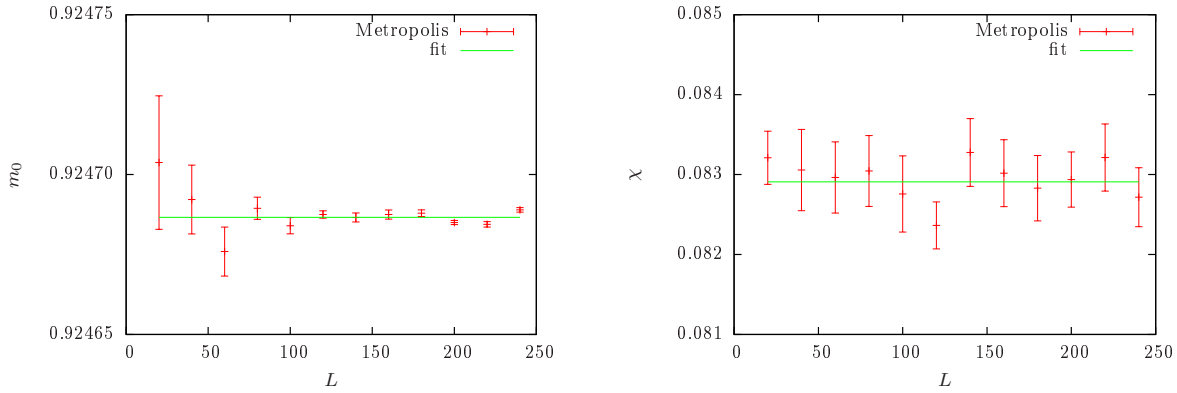

Figure 2.9.: Spontaneous magnetization (left) and magnetic susceptibility (right) from Metropolis simulations of the three-dimensional Ising model at $T^{\mathrm{I}} = 3.2$.

a multimagnetic simulation. This takes advantage of the idea that close to the droplet-strip transition, the complete particle excess is in the largest droplet, i.e., $V_{\mathrm{D}} \approx \frac{\delta N}{1-2\rho_0}$ for a lattice gas with particle-hole symmetry $\rho_{\mathrm{L}} = 1 - \rho_0$. Inserting the definition of the particle excess and taking into account the equivalence to the Ising model, $m = 1 - 2\rho$, this may be written as $V_{\mathrm{D}} \approx \frac{1}{2}(1 - m/m_0)V$. In this case, only the droplet is contributing to the free energy, which leads to the following probability distribution

$$\begin{aligned} P_d(m) &\propto \exp\left\{-\beta\tau_{\mathrm{W}}(V_{\mathrm{D}})^{2/3}\right\} \\ &\approx \exp\left\{-\beta\tau_{\mathrm{W}}\left(\frac{1}{2}\left(1 - m/m_0\right)V\right)^{2/3}\right\}. \end{aligned} \tag{2.32}$$

For $T = 3.2$, a multimagnetic simulation in the regime $m \in [0.63, 0.75]$ suffices to compute the probability distribution in the corresponding range from the final histogram $H(m)$ and the weight function $W(m)$ as $P_d(m) = \frac{H(m)}{W(m)}\left(\sum \frac{H(m)}{W(m)}\right)^{-1}$. Then, plotting $\ln P_d$ versus $-\beta\left(\frac{1}{2}\left(1 - m/m_0\right)V\right)^{2/3}$ yields τ_{W} from a linear fit. For convenience, I plot the size-independent variable $-\beta\left(\frac{1}{2}\left(1 - m/m_0\right)\right)^{2/3}$ and obtained $\tau_{\mathrm{W}}^{\mathrm{Is}}L^2$ instead, using a size-independent fit range. In order to estimate the error, I performed 16 different parallel multimagnetic simulations for each system size and always did the fit within the same size-independent range. Assuming a finite-size scaling dependence

$$\tau_{\mathrm{W}}(L) = \tau_{\mathrm{W}} + aL^{-4/3}, \tag{2.33}$$

allows to extrapolate the infinite-system value for $\tau_{\mathrm{W}}^{\mathrm{Is}}$, see Fig. 2.10. The result of the finite-size scaling is

$$\tau_{\mathrm{W}}^{\mathrm{Is}} = 5.555(1) \qquad \text{for} \quad T^{\mathrm{Is}} = 3.2, \tag{2.34}$$

listed with the other parameters in Table 2.1

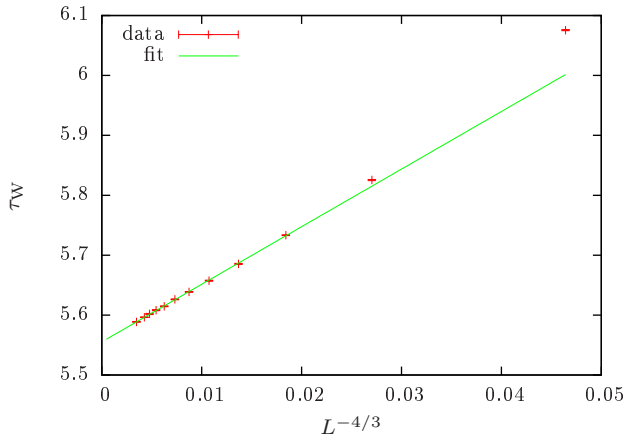

Figure 2.10.: Finite-size scaling of the surface free-energy per unit volume for a temperature $T^{\mathrm{Is}} = 3.2$.

This ansatz may be compared to the surface free-energy per unit volume under the assumption of a spherical shape. The interface tension may be naively considered as the simple arithmetic mean of the anisotropic planar interface tensions in the 3D Ising model for $\beta^{\mathrm{Is}}_{\mathrm{ref}} = 0.31$ [40]:

$$\sigma_{100} = 1.106; \quad \sigma_{110} = 1.127; \quad \sigma_{111} = 1.1324. \tag{2.35}$$

The surface free-energy is then obtained by multiplying the surface of a sphere with volume 1, i.e., $R_1 = \left(\frac{3}{4\pi}\right)^{1/3}$, by the average surface interface tension $\sigma = \frac{1}{3}\left(\sigma_{100} + \sigma_{110} + \sigma_{111}\right) = 1.1218$, yielding:

$$\tau^{\mathrm{Is}}_{\mathrm{W}} = 4\pi R_1^2 \sigma = (36\pi)^{1/3}\sigma \approx 5.425. \tag{2.36}$$

Despite the simplified assumptions, there is a reasonably small difference of only 2% which may be even partially attributed by the small difference in temperature, $T^{\mathrm{Is}} = 3.2$ versus $T^{\mathrm{Is}}_{\mathrm{ref}} \approx 3.23$.

Having collected all necessary parameters for two selected temperatures allows to compare numerical data in the desired Δ-representation of a lattice gas and to compare to the leading-order analytical solution. In Fig. 2.11, the analytical solution $\lambda(\Delta)$ is obtained by numerical minimization of Eq. (2.20) and each data point is again obtained from a multicanonical simulation for either $T_1 = 0.5$ (left) or $T_2 = 0.8$ (right) using the following conversion with parameters

Table 2.1.: Parameters for the Ising model on a cubic lattice at fixed temperature. m_0 and χ are obtained from Metropolis simulations of various lattice sizes (see Fig. 2.9). τ_{W} is obtained from low-temperature series expansion ($T < T_{\mathrm{R}}$) or from a fit to the probability distribution $P(m)$.

T^{Is}	m_0	χ	$\tau^{\mathrm{Is}}_{\mathrm{W}}$
2.0	0.994521(1)	0.00608(1)	≈ 11.443
3.2	0.924687(1)	0.0829(1)	5.555(1)

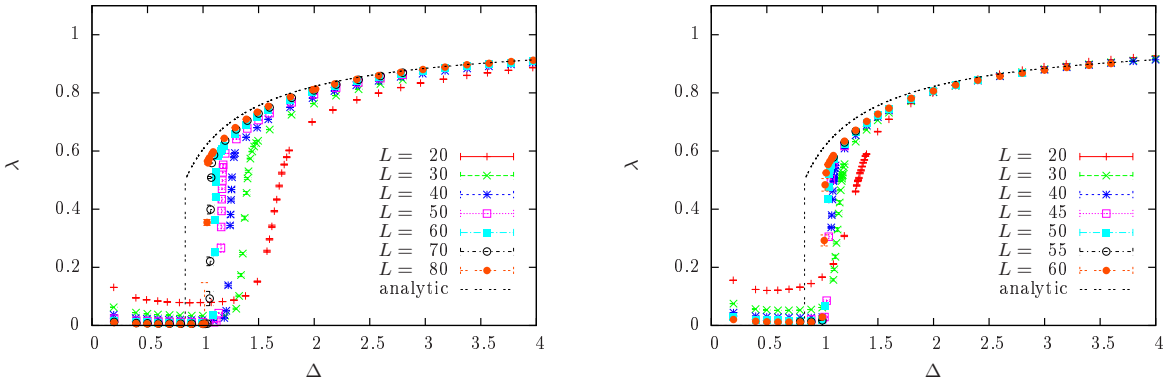

Figure 2.11.: Condensation of a lattice gas in three dimensions at two temperatures around the roughening temperature T_{R}: $T = 0.5 < T_{\mathrm{R}}$ (left) and $T = 0.8 > T_{\mathrm{R}}$ (right). The numerical data from multicanonical simulations show strong finite-size corrections compared to the analytical solution (dashed line).

from Table 2.1, of course translated to the formulation of a lattice gas,

$$\begin{aligned} v_{\mathrm{L}} &= \frac{N - \rho_0 L^3}{1 - 2\rho_0} \\ \lambda &= N_{\mathrm{D}}^{\mathrm{flood}} / v_{\mathrm{L}} \\ \Delta &= \frac{(1-2\rho_0)^2}{2\chi\tau_{\mathrm{W}}} \frac{v_{\mathrm{L}}^{4/3}}{L^3} \end{aligned} \tag{2.37}$$

For consistency with the two-dimensional case, I again consider the largest droplet as the largest connected cluster plus the enclosed holes (determined by a flood-fill algorithm). For the three-dimensional case, this shows only very small deviations in the estimators of both N_{D} and λ. The consideration of these types of geometric clusters is a safe choice for dilute systems. Note that for rather dense lattice systems also the stochastic Swendsen-Wang cluster definition has been considered [41].

With increasing temperature, the density increases drastically so that the largest simulations for $T = 0.8$ includes $N \leq 17183$ particles in $V = 60^3$, compared to $N \leq 5753$ in $V = 80^3$ for $T = 0.5$. The analytical curve depicted is the numerical solution of the analytical result Eq. 2.20 with the parameters from Table 2.1. Both plots show that the qualitative behavior is recaptured as expected. For small Δ only small clusters form, while for large Δ, above the predicted Δ_{c}, a macroscopic fraction of excess is found in a single largest droplet. In contrast to the rather symmetric 2D case, the finite-size transition "density", or the onset of a macroscopic droplet, is shifted to larger $\Delta_{\mathrm{c}}(L)$ well above Δ_{c}^{3D} for both temperatures considered. However, it may be argued that the finite-size transition points shift towards the predicted leading-order value Δ_{c}.

To study this behavior in more detail, consider a horizontal line at $\lambda_c = 1/2$. For each system size, I performed additional simulations in this vicinity, estimating the crossing of the finite-size estimator $\lambda_L(N)$ with that horizontal line. I define N_{c} as the total number of particles at fixed volume and temperature for which $\lambda_L(N_{\mathrm{c}}) \approx 1/2$. In three dimensions, the proposed leading-order correction for $\rho_{\mathrm{c}} = N_{\mathrm{c}}/V$ is given in Eq. (2.24). Further, I assume empirical

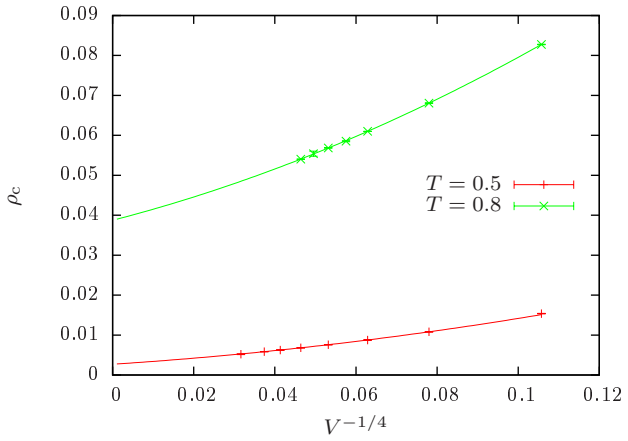

Figure 2.12.: Finite-size scaling test in three dimensions including empirical higher order corrections to the predicted leading-order one.

higher order corrections as polynomial powers of the leading order. Figure 2.12 shows this finite-size scaling ansatz with $\rho_c = a + bV^{-1/4} + cV^{-2/4}$ for the two temperatures $T_1 = 0.5$ and $T_2 = 0.8$. Both fits have appropriate $\chi^2(0.5) \approx 0.3$ and $\chi^2(0.8) \approx 0.4$. The finite-size extrapolation yields $a(0.5) = 0.00269(3)$ and $a(0.8) = 0.0387(3)$. This should be compared to the low-temperature series expansion and Metropolis Ising simulations listed in Table 2.1, namely $\rho_0(0.5) = 0.002739(1)$ and $\rho_0(0.8) = 0.03765(1)$. The result is astonishingly similar given rather rough fit errors and the fact that this is only an empirical higher-order correction neglecting additional sources of corrections. In addition, the fit parameters yield an estimate of the rescaled condensation density Δ_c from Eq. (2.24), namely in 3D

$$\Delta_c^{3D} \approx \frac{(1-2\rho_0)^{2/3}}{2\hat{\kappa}\tau_W} b^{4/3}. \tag{2.38}$$

With the fit parameters $b(0.5) = 0.066(1)$ and $b(0.8) = 0.263(9)$ and the self-consistent use of the fit parameter a for ρ_0 this yields $\Delta_c(0.5) \approx 0.76(2)$ and $\Delta_c(0.8) \approx 0.69(3)$. This is below the predicted value $\Delta_c^{3D} = 0.8399...$ with multiple sources of errors and systematic uncertainties. The error given is obtained from error propagation of the corresponding fit errors from all contributing parameters. Together with Fig. 2.11, however, it may be conjectured that also in three dimensions the analytical prediction from Biskup et al. [10] should remain valid for large system sizes.

Of course, the transition density is only a rough estimate with a lot of computational overhead from the multicanonical simulation, as this was not the focus of my attention. In order to measure the finite-size corrections precisely, it would be beneficial to consider the derivative of $N_D(\rho)$, which could be obtained by either grand-canonical simulations or a suitable set of fixed-density Metropolis simulations (or micromagnetic Ising simulations [15–17]). This, however, has the systematic disadvantage that for finite lattice systems ρ is not continuous. The following section shows a different (orthogonal) approach with the continuous variable T, where the multicanonical method is naturally applicable.

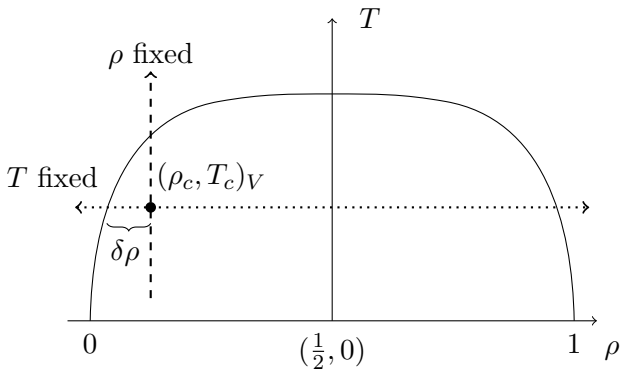

Figure 2.13.: Sketch of the infinite system-size transition (solid line) together with the finite-size scaling directions in either the density (T fixed) or temperature (ρ fixed). The transition line may be understood as $T_0(\rho)$ or similarly $\rho_0(T)$. At the crossing point of both schemes a finite system of size V may be constructed for which $(\rho, T) = (\rho_c, T_c)_V$ corresponds to the finite-size condensation transition. Adapted from Ref. [21].

2.3. Thinking orthogonal: crossing the phase boundary at fixed density

In the previous discussion, the liquid-vapor system was considered at fixed temperature. In this way, free energies may be approximated by density fluctuations and surface contributions using infinite-size canonical parameters (at constant temperature) which do not change with system size. This finite-size scaling scheme is depicted in Fig. 2.13 by the horizontal dashed line, on which for finite systems the transition density may be found. For the infinite system (or the thermodynamic limit) the transition point is recovered as the intersection of the solid and the dashed line.

Alternatively, one may consider the "orthogonal" scheme: keeping the density fixed leads to a shift in the transition temperature as shown by the vertical dashed line in Fig. 2.13. More explicitly, consider a sufficiently dilute liquid-vapor system of N particles in a (periodic) box of volume V, where the condensation/evaporation transition temperature T_c separates a homogeneous supersaturated gas phase from a mixed phase of a droplet in equilibrium with surrounding bulk gas. It needs to be mentioned that both schemes are working in the canonical ensemble for each point (ρ, T). So in fact, any finite-size transition point $(\rho_\mathrm{c}, T_\mathrm{c})_V$ belongs to one fixed-T and one fixed-ρ scheme, simultaneously. The same holds for any canonical function $f(\rho, T)$. Thus, the orthogonal crossing schemes are equivalent and we may translate a functional dependence such as $f(\rho, T)V^\alpha = 1$ from one scheme to another by a Taylor series expansion. Then, expanding around some T^* yields

$$V^{-\alpha} = f(\rho, T^*) + f'(\rho, T^*)(T - T^*) + ..., \tag{2.39}$$

which may be solved for T. The remaining task is to identify and evaluate suitable functional dependencies. This section is in large parts content of Ref. [21].

Using this procedure, I will discuss for a fixed density $\rho = N/V$ the finite-size correction to the transition temperature $T_c - T_0 \propto N^{-1/(d+1)}$ and the scaling of the finite-size rounding (the

temperature width of the finite-system transition region) $\Delta T = T - T_c \propto N^{-d/(d+1)}$. I will identify that these leading-order corrections are related to the linear extension of the droplet at coexistence $R \propto N^{1/(d+1)}$, which thus seems to become the relevant length scale for large system sizes.

In the following, I will consider partially rigorous results in order to discuss the scaling for intermediate sized systems and in the limit of large systems. Each discussion of finite-size scaling corrections is directly accompanied by numerical results from parallel multicanonical simulations (see Sec. 5.4). Working in the fixed-ρ scheme unleashes the full power of the multicanonical simulation, recovering the observables as full (precise) functions of the temperature. All simulations in this section are performed at constant density $\rho = 10^{-2}$. If not stated differently, the numerical results are obtained from parallel multicanonical simulations with $p = 64 - 128$ cores and 2.56 million measurements in total. Initial simulations of small systems with parallel tempering pre-runs were used to estimate the multicanonical energy range, which was adjusted for larger system sizes such that the condensation transition was sampled. Relevant data presented in the plots are listed in Tables in the appendix.

2.3.1. Finite-size scaling of the condensation/evaporation transition temperature

Recalling Sec. 2.2, Biskup et al. [10] showed a vanishing probability of intermediated-sized droplets for a d-dimensional liquid-vapor system at fixed temperature. Moreover, they calculated the fraction of particle excess in the largest droplet λ as a "universal" function of a rescaled density Δ (see Eq. (2.19)). The resulting leading-order transition density for finite systems Eq. (2.24) is in good agreement with additional results in the literature [12, 13]. Considering now a *fixed density* ρ the involved infinite-size constants become functions of the temperature, namely the background densities $\rho_i(T)$, the reduced isothermal compressibility $\hat{\kappa}(T)$, the surface free energy of a (Wulff shaped) droplet of unit volume $\tau_{\mathrm{W}}(T)$ and the particle excess $\delta N = (\rho - \rho_0(T))\, V$. In the spirit of the aforementioned Taylor expansion, the leading-order result in Eq. (2.19) may be written as

$$\Delta^{\frac{d}{d+1}} V^{-\frac{1}{d+1}} = f(\rho, T) = \frac{\rho - \rho_0(T)}{\rho_{\mathrm{L}}(T) - \rho_0(T)} \left(\frac{(\rho_{\mathrm{L}}(T) - \rho_0(T))^2}{2\hat{\kappa}(T)\tau_{\mathrm{W}}(T)} \right)^{\frac{d}{d+1}}. \tag{2.40}$$

At the condensation transition $\Delta = \Delta_{\mathrm{c}}$ is constant and the left hand side of Eq. (2.40) is depending only on the system size V. Then, for a fixed finite system size, a suitable combination of T and ρ solves Eq. (2.40) yielding the finite-size condensation transition point at $(\rho, T) = (\rho_{\mathrm{c}}, T_{\mathrm{c}})_V$ in Fig. 2.13. This transition point may be obtained either numerical exact or by a Taylor expansion. In the latter case, keeping $\rho = N/V$ constant, one proceeds by expanding Eq. (2.40) around the infinite-system transition temperature T_0. Then $f(\rho, T) = f(\rho, T_0) + f'(\rho, T_0)(T - T_0) + \ldots$, where the first term vanishes due to $\rho_0(T_0) = \rho$. Solving this for the finite-size condensation transition temperature $T = T_{\mathrm{c}}$ yields

$$T_{\mathrm{c}} = T_0 + \frac{\Delta_{\mathrm{c}}^{\frac{d}{d+1}}}{f'(\rho, T_0)} V^{-\frac{1}{d+1}} + \mathcal{O}\left(V^{-\frac{2}{d+1}}\right). \tag{2.41}$$

In terms of the number of particles this means to first order

$$T_{\mathrm{c}} - T_0 \propto N^{-1/(d+1)}. \tag{2.42}$$

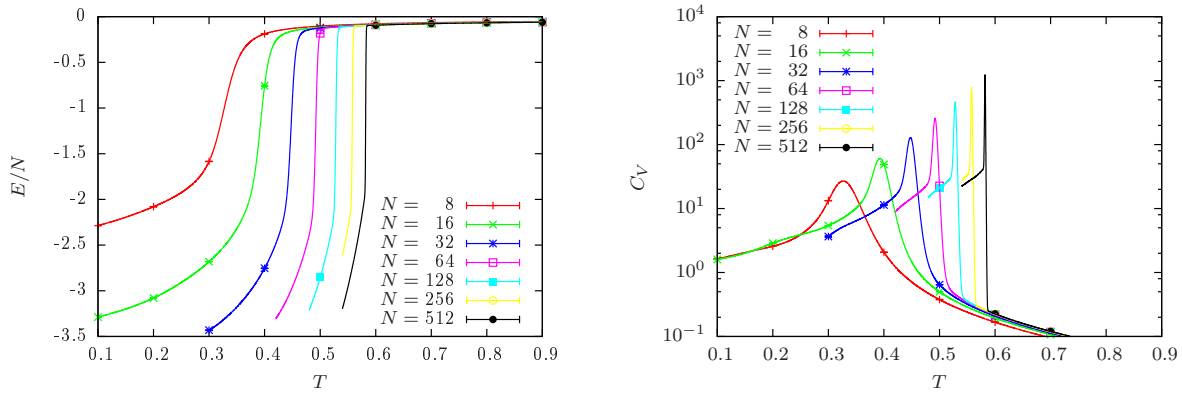

Figure 2.14.: The canonical average energy (left) and the specific heat (right) of a Lennard-Jones particles system at constant density $\rho = 10^{-2}$.

As mentioned before, already Eq. (2.40) is a leading-order result. For more details see the discussion and numerical results for empirical higher order corrections in Sec. 2.2.

The equivalence of the lattice gas and the Ising model allow to explicitly calculate the finite-size corrections. Recalling Eq.(2.25) this yields

$$f(m, T^{\mathrm{Is}}) = \frac{1}{2}\left(1 - \frac{m}{m_0(T^{\mathrm{Is}})}\right)\left(\frac{2m_0(T^{\mathrm{Is}})^2}{\chi(T^{\mathrm{Is}})\tau_{\mathrm{W}}^{\mathrm{Is}}(T^{\mathrm{Is}})}\right)^{\frac{d}{d+1}}. \tag{2.43}$$

In two dimensions, the involved quantities are known analytically or up to arbitrary precision, see Sec. 2.1.1. This allows to numerically evaluate Eq. (2.43) in Eq. (2.40), fixing the density as well as the volume and solving with a bisection algorithm for the corresponding transition temperature $T_{\mathrm{c}} = T_{\mathrm{c}}^{\mathrm{Is}}/4$. When comparing to the simulation results later, I will refer to this as the *full solution* of Eq. (2.40). The infinite-size transition temperature $\tilde{T}_0$ that belongs to a given fixed density is obtained by inversion of the Onsager solution and, thus, is exact for the two-dimensional lattice gas. In three dimensions, I make use of low-temperature series expansions of the spontaneous magnetization (see Sec. 2.1.1). Solving $\rho - \rho_0(\tilde{T}_0) = 0$ allows to estimate the infinite-system transition temperature $\tilde{T}_0$. In the following this will be referred to as the solution from *low-temperature series expansion*.

Numerical results: In order to identify the transition temperature at fixed density, I consider two (related) observables and their thermal derivatives for various system sizes: the average energy and the fraction η of particles in the largest droplet. For the 3D Lennard-Jones gas see Fig. 2.14 and Fig. 2.15, respectively. The canonical data points are obtained from time-series reweighting and the connected lines from histogram reweighting of parallel multicanonical simulations. The condensation transition temperature is now identified as the temperature where the change in observable is maximal, i.e., where the thermal derivative shows an extremum. Numerically, this was achieved by finding the zero-crossings of the successive thermal derivative, again applying the jackknife error analysis.

Figure 2.14 clearly shows the first-order nature of the transition, where an entropy-dominated phase at high temperature (with almost equal average energy) and an energy dominated phase at low temperature are connected by a jump in the energy. The specific heat $C_V =$

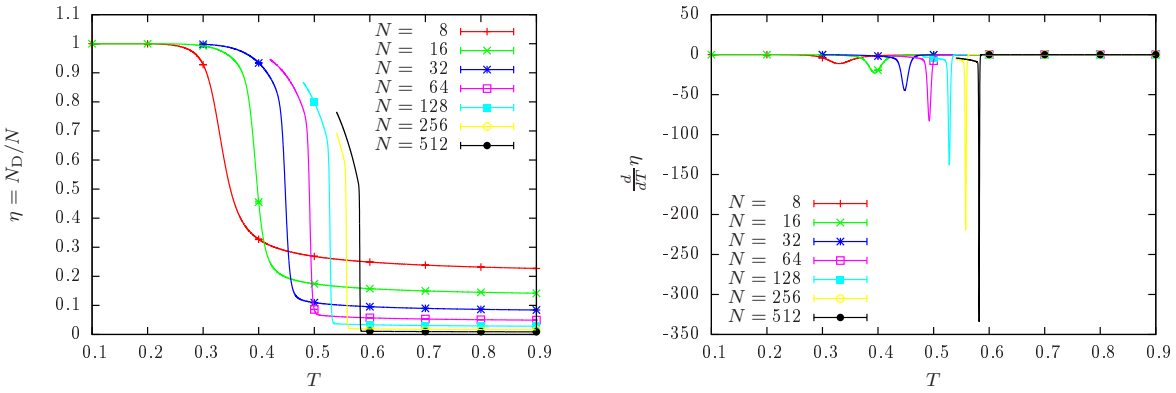

Figure 2.15.: The fraction of particles in the largest droplet η (left) and it's thermal derivative (right) of a Lennard-Jones particles system at constant density $\rho = 10^{-2}$.

$\beta^2 \left(\langle E^2\rangle - \langle E\rangle^2\right)/N$ shows a pronounced peak that increases in height and decreases in width with increasing system size. The latter will be discussed in more detail in Sec. 2.3.2. At the condensation transition, a single macroscopic droplet forms in equilibrium, which can be seen in Fig. 2.15 by a sudden increase in the fraction η of particles in the largest droplet. For small systems a large finite-size effect is evident in the high-temperature range, where always at least one particle is in the largest droplet by definition and thus η is non-zero. Below the condensation transition, the droplet grows with decreasing temperature in a similar qualitative way as with increasing supersaturation in Sec. 2.2. Again, the condensation transition is accompanied by a sharp peak in the thermal derivative.

Numerical results for the scaling of the estimated transition temperature are shown in Fig. 2.16 for the 2D lattice gas and in Fig. 2.17 for the 3D lattice and Lennard-Jones gas. As described above, the transition temperatures are estimated from the largest peak of the specific heat (red pluses) and the smallest minimum of the thermal derivative of the fraction of particles in the largest cluster $\frac{d\eta}{dT}$ (green crosses). Both estimates are remarkably similar as expected for first-order phase transitions, and hence the plot only shows local fits to the estimated transition temperature derived from the specific heat. All three cases show for large system sizes the proposed finite-size scaling behavior from Eq. (2.42), shown by a good fit quality (a reduced χ^2 per degree of freedom of about 1). The leading-order fit (dashed dark blue fit) requires rather large system sizes in accordance with the literature. Including empirical higher-order corrections of the form

$$T_c = T_0 + aN^{-1/(d+1)} + bN^{-2/(d+1)} \tag{2.44}$$

allows to extend the fit range to cover also the intermediate-sized systems (dotted light blue fit). Considering only the intermediate system sizes allows to identify a different effective scaling behavior. In this case, the transition temperature may be locally described by a $N^{-1/d}$ behavior (dashed-dotted orange fit), which may be justified by a sufficiently good χ^2. The scaling of the intermediate regime is consistent with observations for flexible polymer aggregation (see Sec. 3.4 and Ref. [42]), where a large fraction of monomers is involved in the formation of the aggregate. This may be justified in terms of the above mentioned dependence on the linear extension of the condensate. If the condensate includes almost all constituents, then the linear

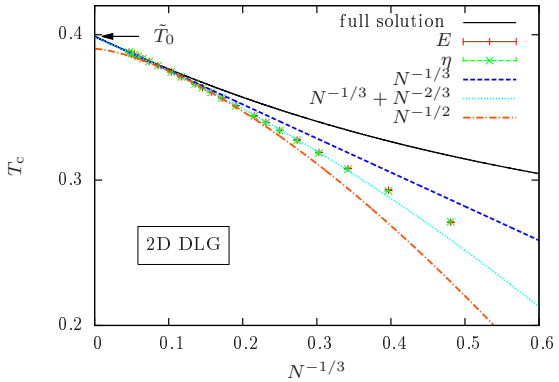

Figure 2.16.: Finite-size scaling of the droplet evaporation/condensation transition temperature in a 2D lattice gas: compared to exact results ($\tilde{T}_0$) and the full solution of Eq. (2.40).

extension of the homogenous, isotropic condensate is just $N^{1/d}$. For details of the individual fits see Table 2.2 and the following discussion.

The finite-size scaling of the 2D lattice gas, shown in Fig. 2.16, may be compared directly to the exact analytical solution of the equivalent Ising model for an infinite system and to the full solution of Eq. (2.40) using Eq. (2.43) for a finite system. As expected, the full solution shows large deviations for small system sizes but starts to describe the finite-size scaling approximately with increasing system size. The exact analytical result is obtained by inverting Onsager's solution of the magnetization $m_0(\tilde{T}_0) = 0.98$ (see Eq. (2.5)) with $T = T^{\mathrm{Is}}/4$. This yields $\tilde{T}_0 = 0.39882$ and is shown in Fig. 2.16 by the arrow. The leading-order fit for the largest system sizes $N \geq 2500$ yields the infinite-size transition temperature $T_0 = 0.39884(3)$ with $\chi^2 \approx 0.4$. This is in good agreement with the analytical result. Including the empirical next order and fitting $N \geq 400$ yields $T_0 = 0.3982(1)$ with $\chi^2 \approx 1.8$, which deviates slightly from the exact result. This may be taken as a hint that the empirical higher-order term is merely an effective correction and additional corrections of the same order are apparent. Having the full solution of Eq. (2.40) allows in addition to compare the amplitude of the leading-order correction $aN^{-1/3}$. To this end, consider the power-series expansion of the full solution Eq. (2.41). Then, Eq. (2.43) may be numerically differentiated making use of the analytical solution for m_0, the series expansion for χ up to 300^{th} order and the integral solution of τ_{W} described in Sec. 2.1.1. Converted back to lattice gas units this gives $a = \Delta_c^{2/3}\rho^{1/3}/4f'(m, T_0^{\mathrm{Is}}) \approx -0.239$. On the other hand, the leading-order and higher-order fits yield $a = -0.234(1)$ and $a = -0.214(2)$ respectively. Again, the leading-order fit for large N is in good agreement with the analytical (approximate) result, while the higher-order fit covers a wider range of numerical data.

For the two-dimensional lattice gas the intermediate scaling regime is not very prominent. A least-square fit to $N^{-1/2}$ in the (already small) range $N = [324, 900]$ still yields a $\chi^2 \approx 8.0$. Moreover, the infinite-size extrapolation is clearly wrong. However, the intermediate scaling regime in polymer aggregation [42] was observed in three dimensions which suggests that the prominence of this regime may depend on the dimension.

The finite-size transition temperatures for the three-dimensional lattice gas and Lennard-Jones gas are shown in Fig. 2.17. The infinite-size transition temperature for the 3D lattice gas

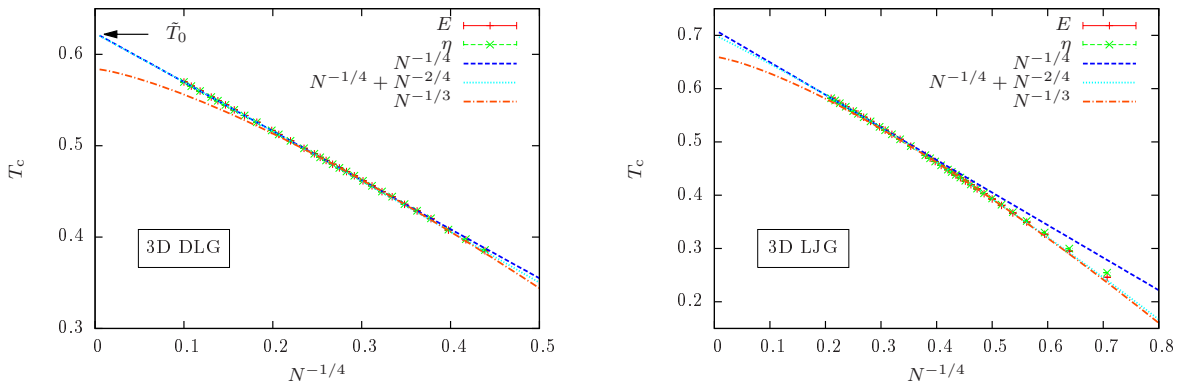

Figure 2.17.: Finite-size scaling of the droplet evaporation/condensation transition temperature for the 3D lattice (left) and Lennard-Jones gas (right). The infinite-system extrapolation of the lattice gas is compared to low-temperature series expansions yielding $\tilde{T}_0$.

may be compared to the inversion of low-temperature series expansions, see Eq. (2.8), which leads to $\tilde{T}_0 \approx 0.622$ shown in the figure by the arrow. The leading-order fit for the 3D lattice gas with $N \geq 2160$ yields $T_0 = 0.62341(4)$ with $\chi^2 \approx 0.6$ which is in the vicinity of the (not exact) low-temperature expansion result. Including the empirical higher-order term $N^{-2/4}$ still shows a similar fit with a very small contribution from the higher order term. Considering, however, only an intermediate regime allows to fit the $N^{-1/3}$ behavior with qualitatively good local agreement.

In the case of the Lennard-Jones gas one may best see the arising peculiarities. The leading-order fit for $N \geq 160$ yields $T_0 = 0.7106(4)$ with $\chi^2 \approx 0.8$ but shows clear deviations for small system sizes. Including the empirical next order for $N \geq 10$ yields $T_0 = 0.7011(4)$ with $\chi^2 \approx 1.1$ in rough agreement with the leading-order result, and recaptures the deviation of the small system sizes. This is consistent with results for the same Lennard-Jones model [43, 44]. However, considering an intermediate regime with the Ansatz $N^{-1/3}$ yields a good fit with $T_0 = 0.6597(4)$ and $\chi^2 \approx 1.0$, which deviates strongly from the $N^{-1/4}$ fit. Another interesting observation is the relatively small maximal system size of $N = 512$ in a box of length $L \approx 37.1 \approx 41.7\sigma$ necessary to qualitatively recover the predicted scaling behavior. Other studies [19], following the fixed-T scheme, considered system sizes up to $L = 100\sigma$ at $T \approx 0.68$ (for their parameterization) with typical particle numbers $N \approx 15\,800$. Their data for the critical density versus linear system size L still showed a strong curvature with effective exponents smaller than -0.89, while extrapolation of these effective exponents recovered the theoretical predication $L^{-0.75}$. A direct fit of the leading-order power-law exponent to the present data yields $N^{-0.28(1)} \propto L^{-0.84(2)}$ already for remarkably smaller system sizes. This implies that an orthogonal phase boundary crossing may lead in certain situations to reduced finite-size corrections and serves as a useful, complementary approach.

If the largest system sizes were not present, the intermediate regime could be, at least in three dimensions, easily interpreted as the leading-order finite-size scaling corrections, especially if no reference temperature is available. Including larger system sizes, it locally appears to be an effective intermediate scaling regime. It is, however, covered by the theoretically predicated scaling behavior including empirical higher-order corrections.

Table 2.2.: Results of different fit functions to the condensation/evaporation transition temperature including fit errors [21]. If no upper range is provided, it refers to $N_{\max} = 10\,000$ for the lattice gas cases (2D and 3D) and $N_{\max} = 512$ for the Lennard-Jones gas. For the lattice systems, reference values for the infinite-size condensation temperature $\tilde{T}_0$ are obtained from the Onsager solution (2D) [23, 24] and from low-temperature series expansions (3D) [34, 35].

Fit		2D DLG	3D DLG	3D LJG
$T_0 + aN^{-1/(d+1)}$	Range	[2500:]	[2160:]	[160:]
	T_0	0.39884(3)	0.62341(4)	0.7106(4)
	χ^2	0.4	0.6	0.8
$T_0 + aN^{-1/(d+1)} + bN^{-2/(d+1)}$	Range	[400:]	[1663:]	[10:]
	T_0	0.3982(1)	0.6229(4)	0.7011(4)
	χ^2	1.8	1.9	1.1
$T_0 + aN^{-1/d}$	Range	[324: 900]	[68: 243]	[12: 48]
	T_0	0.3903(2)	0.5840(3)	0.6597(4)
	χ^2	8.0	1.6	1.0
Reference	$\tilde{T}_0$	0.39882...	≈ 0.622	

2.3.2. Finite-size rounding of the condensation/evaporation transition

Analogue to the transition temperature, one may argue that the rounding of the transition at fixed density should scale with the system size in the same way as the rounding of the transition density/magnetization at fixed temperature. Comparing to Eq. (2.30) in Sec. 2.2, one would expect a finite-size rounding proportional to $V^{-d/(d+1)}$.

In a two-state approximation [13], let us reduce the consideration to a condensate and a gas phase with free-energy F_{cond} and F_{gas}, respectively. The partition function becomes

$$Z = e^{-\beta F_{\text{gas}}} + e^{-\beta F_{\text{cond}}} = e^{-\beta F_{\text{gas}}} \left(1 + e^{-\beta \Delta F}\right), \tag{2.45}$$

with $\Delta F = F_{\text{cond}} - F_{\text{gas}}$. From this partition function, the canonical free-energy becomes $\beta F = -\ln Z = \beta F_{\text{gas}} + \ln\left(1 + e^{-\beta \Delta F}\right)$ and the canonical expectation value of the energy is then

$$\langle E \rangle = -\frac{\partial \ln Z}{\partial \beta} = \frac{E_{\text{gas}} + e^{-\beta \Delta F} E_{\text{cond}}}{1 + e^{-\beta \Delta F}}. \tag{2.46}$$

The free-energy difference between both phases $\beta \Delta F$ thus weights the contribution of the phases to the expectation value. At $T = T_{\text{c}}$ both states are equally probable which is reflected in $\beta \Delta F = 0$. Otherwise, the condensate phase is either suppressed or enhanced.

The rounding of the transition may be related to the region where thermal fluctuations are of the order $k_{\text{B}}T$ [13, 18], i.e., $\Delta F \sim k_{\text{B}}T$ or $\beta \Delta F \sim 1$. In order to estimate this, consider the expansion of $f(T) = (\beta F)(T)$ around T_{c} – the temperature at which the probability distribution of this simplified model shows two peaks. With $F = E - TS$ and $\partial / \partial T (\beta F) = -\frac{1}{k_{\text{B}} T^2} E$ this

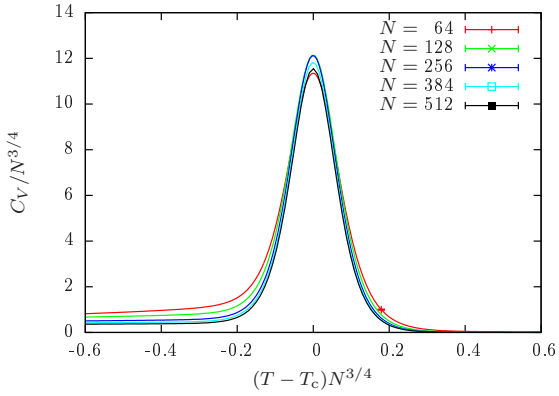

Figure 2.18.: Example of the transition rounding on the rescaled specific heat peak for the 3D Lennard-Jones gas in Fig. 2.14 (right). The axes are rescaled according to the leading-order scaling behavior.

yields for the free-energy difference between the coexisting phases

$$\beta\Delta F = (\beta\Delta F)|_{T_c} - \left(\frac{1}{k_B T^2}\Delta E\right)\Big|_{T_c} (T - T_c) + \dots \,. \tag{2.47}$$

The free-energy difference vanishes at T_c in the limit of large system sizes, considering that both phases contribute with equal probability. In the two-state approximation this holds by definition and the first term vanishes. Moreover, the energy difference $\Delta E = E_{\text{cond}} - E_{\text{gas}}$ may be estimated in terms of the individual phases. For the gas phase, the energy is negligible small because the particles may be considered non-interacting. Thus, the energy difference is dominated by the condensate phase, which is a mixed phase of droplet and gas and thus depends on the droplet volume $\propto R^d$. The volume of the droplet, however, does not grow linearly with system size but in leading order $R \propto N^{1/(d+1)}$ (see below) such that $\Delta E \sim N^{d/(d+1)}$. The finite-size corrections from the condensation temperature appear merely as corrections to the energy difference, such that Eq. (2.47) simplifies to $\beta\Delta F \sim (N^{d/(d+1)}/k_B T_0^2)\ \Delta T$ in the limit of large system sizes. The condition $|\beta\Delta F| \sim 1$ yields to leading order the rounding width

$$\Delta T \propto N^{-d/(d+1)}, \tag{2.48}$$

see also Sec. 2.2 for the d-dimensional result of $\Delta\rho$. Notice that the radius of the droplet R was used as relevant length scale such that the rounding may be identified, for large system sizes, as $\Delta T = T - T_c \propto R^{-d}$.

Numerical results: For the discussion of the finite-size rounding I focus on the specific heat only. Figure 2.18 shows the rescaled specific heat from Fig. 2.14 for the 3D Lennard-Jones gas. The axes are rescaled according to the theoretical leading-order scaling behavior, where the peak of the specific heat should scale inverse proportional to the rounding in the temperature for large system sizes [39]. This may be observed qualitatively in the figure, as well as the almost constant width of the peak. However, small deviations already indicate that additional corrections are apparent.

The rounding of the transition is estimated as the half-width of the specific heat peak, defined

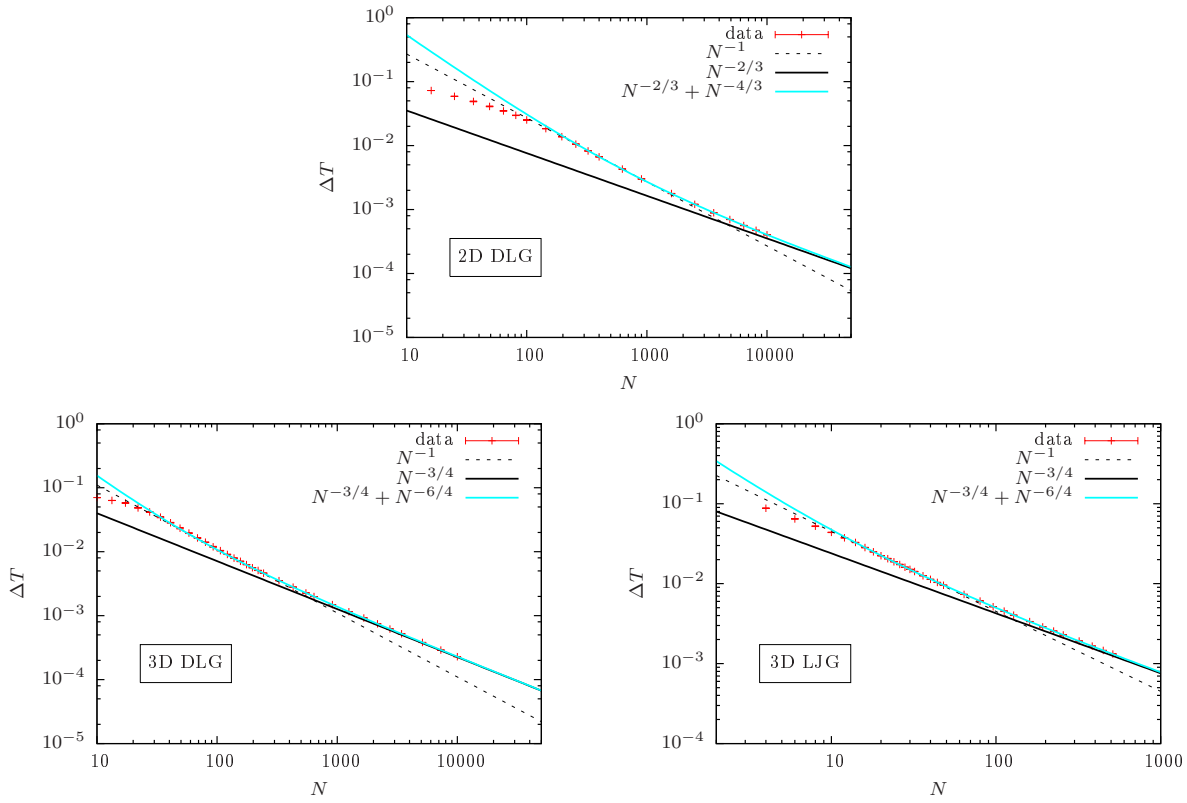

Figure 2.19.: Finite-size rounding of the droplet evaporation/condensation transition for the 2D lattice gas (top), 3D lattice gas (bottom, left) and Lennard-Jones gas (bottom, right).

as the width where $C_V \geq \frac{1}{2} C_V^{\max}$. Errors are obtained by jackknife error analysis. Figure 2.19 shows the finite-size rounding of the condensation/evaporation transition for all three considered systems. Two different scaling regimes may be clearly identified: for intermediate system sizes the data show a $\Delta T \propto N^{-1}$ behavior (dashed line) and for large system sizes one may observe the predicted $\Delta T \propto N^{-d/(d+1)}$ behavior (solid line). Again, having only data available in the intermediate regime would support a wrong finite-size scaling behavior consistent with the intermediate regime for the transition temperature. As for the transition temperature, the intermediate regime may be covered by the theoretically predicted scaling if empirical higher-order corrections are included (light blue fit), i.e., $\Delta T = a' N^{-d/(d+1)} + b' N^{-2d/(d+1)}$.

For the 2D lattice gas, Fig. 2.19 (top), the smallest systems show no systematic behavior which accounts for the strong deviations in the case of the transition temperature, recall Fig. 2.16. However, for the rounding there is also in two dimensions a clear intermediate scaling regime observable, where one can see a direct particle or volume dependence (ρ fixed). The onset of the large-system regime is in good agreement with the fit range in the transition temperature necessary to recapture the leading-order scaling behavior. Considering the empirical higher-order corrections yields a $\chi^2 \approx 0.7$ including already system sizes $N \geq 324$ and thus including the intermediate regime.

In the case of the 3D lattice and Lennard-Jones gas, Fig. 2.19 (bottom, left and right), the intermediate regime is already apparent for quite small system sizes. The N-dependence up to

the crossover to the large-system regime is clearly visible. Also, the location of the crossover is again consistent with good choices of leading-order finite-size scaling fit ranges for the transition temperature. A fit to the transition rounding including the empirical higher-order corrections allows in both cases to partially cover the intermediate regime. For the lattice gas the fit yields $\chi^2 \approx 2.3$ with $N \geq 175$; and for the Lennard-Jones gas, the fit yields $\chi^2 \approx 0.9$ with $N \geq 16$. In all cases, the exponents were kept fixed.

The present results clearly confirm the "large"-system scaling behavior. This should be compared to previous studies of the lattice gas in two and three dimensions at fixed density, which showed significant deviations from the predicted exponents for the transition rounding [20], using average densities of states from Wang-Landau simulations. In 3D they find an effective scaling of the transition rounding with $L^{-2.45(2)} \propto N^{0.82(1)}$. A direct fit of a power-law behavior to the present 3D lattice gas data yields the effective exponents $N^{-0.78(1)}$, $N^{-0.77(1)}$ and $N^{-0.76(1)}$ for $N \geq 1663$, $N \geq 3430$ and $N \geq 5120$ respectively. This is close to the predicted scaling $N^{-0.75}$ and seems to systematically approach it. Thus, the present data support the predicted scaling of the transition rounding in the condensation/evaporation transition both qualitatively and quantitatively.

The finite-size rounding seems to be a good observable to identify the (effective) intermediate scaling regime. Theoretical reasoning relates the width of the transition to the fluctuations in the system [39], which should depend on the inverse volume of the relevant system size. One may argue now, that the relevant system size for the condensation/evaporation transition is the transition droplet itself, which will be discussed in the following subsection.

2.3.3. Finite-size scaling of the droplet size at condensation/evaporation

The size of the droplet in equilibrium with surrounding vapor directly at the condensation/evaporation transition, was shown to have a non-trivial dependence on the system size, $R \sim L^{d/(d+1)}$, already in the 80s [14]. The leading-order analytical results may be used in order to recapture this scaling at the condensation transition where the fraction λ_c of particle excess will be in the largest droplet. It follows for the droplet volume $V_\mathrm{D} = (\rho_\mathrm{L} - \rho_0)^{-1} \lambda_\mathrm{c} \delta N_\mathrm{c}$ [10]. The volume of an ideal droplet may be expressed in general by $V_\mathrm{D} = S_d R^d$, where S_d is a geometric shape factor that allows to describe both spherical and cubic droplets, where the latter may occur in lattice systems below the roughening transition [17]. Comparing both equations for the droplet volume, inserting δN_c from Eq. (2.19), and solving for the radius yields

$$R = \left(S_d^{-1} V_\mathrm{D}\right)^{1/d} \sim V^{\frac{1}{d+1}}. \tag{2.49}$$

Again, for a fixed density this leads to $R \sim N^{1/(d+1)}$. Moreover, the leading finite-size scaling corrections in Eq. (2.42) and Eq. (2.48) may be expressed in powers of R^{-1}, i.e., $T_\mathrm{c} - T_0 \propto R^{-1}$ and $\Delta T \propto R^{-d}$, respectively.

In practice, the radius – or more generally the shape – of an object is difficult to measure. However, the scaling of the droplet radius is equivalent to the scaling of the droplet size $V_\mathrm{D} \sim N^{d/(d+1)}$. For uniform particles, this is again equivalent to the droplet "mass" or number of particles in the droplet $N_\mathrm{D} \sim N^{d/(d+1)}$. Then, the fraction $\eta_\mathrm{D} = N_\mathrm{D}/N$ of particles in the transition droplet should scale as $\eta_\mathrm{D} \sim N^{-1/(d+1)}$.

Numerical results: At the condensation transition, the reweighted multicanonical expectation values yield the canonical average from equal contributions of both the droplet and the gas

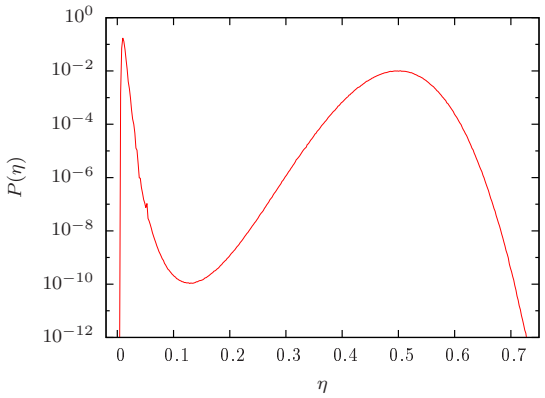

Figure 2.20.: Probability distribution of the fraction η of particles in the largest droplet for the Lennard-Jones gas with $N = 512$ total particles.

phase. Due to the strong finite-size effects in the gas phase ($\eta > 1/N$ always) this average does not allow to estimate the size of the droplet at coexistence. However, time-series reweighting of the multicanonical data allows to calculate the full droplet-size distribution $P(\eta)$ at the condensation transition T_c. This is achieved by going through the time series and adding the reweighting factor to the corresponding bin of the distribution $P(N_\mathrm{D})$ of particles in the largest cluster, which is discrete by definition. The desired distribution is then obtained by normalizing the x-axis with the number of particles N. An example is shown in Fig. 2.20 where one may clearly see two double peaks: one at small fractions corresponding to separated particles and one with larger fractions connected to a single macroscopic droplet plus surrounding, disconnected particles. Now, from this distribution one may estimate the expectation value of the fraction η of particles in the largest droplet inside the droplet phase, i.e. in the right peak of the figure. The distribution is properly normalized, such that $\int_0^1 P(\eta)d\eta = 1$. Assuming that, at coexistence, both phases are at equal weight, I identify η_min such that $\tilde{Z} = \int_{\eta_\mathrm{min}} P(\eta)d\eta = 0.5$. Then

$$\langle\tilde{\eta}\rangle_{T_\mathrm{c}} = \frac{1}{\tilde{Z}} \int_{\eta_\mathrm{min}} \eta P(\eta)|_{T_\mathrm{c}}\, d\eta \tag{2.50}$$

is a robust estimator of the average transition droplet size at coexistence for sufficiently large systems. Moreover, dealing with discrete N_D the integrals become discrete sums. Errors are again estimated using the jackknife error analysis.

I additionally apply standard Metropolis simulations for large system sizes directly at an estimated transition temperature, making use of the finite-size scaling result for the transition temperature in Sec. 2.3.1. In this case, the system gets prepared in the droplet phase, preferably already in an equilibrium droplet with surrounding vapor. Due to the first-order nature of the transition and the resulting barrier, the Metropolis simulation should sample only within the droplet phase for a sufficiently long time. The average transition droplet size is in this case obtained by the statistical average including also an integrated autocorrelation time analysis for the error. This needs a lot less computing time than the parallel multicanonical simulation, however, requires an accurate estimate of the transition temperature. Here, I consider only the leading-order behavior $T_c(N) = T_0 + aN^{-1/(d+1)}$ of the largest system sizes. More explicitly, I consider the estimates $(T_0, a) = (0.623, -0.537)$ for the 3D lattice gas and $(T_0, a) = (0.710, -0.611)$ for the 3D Lennard-Jones gas.

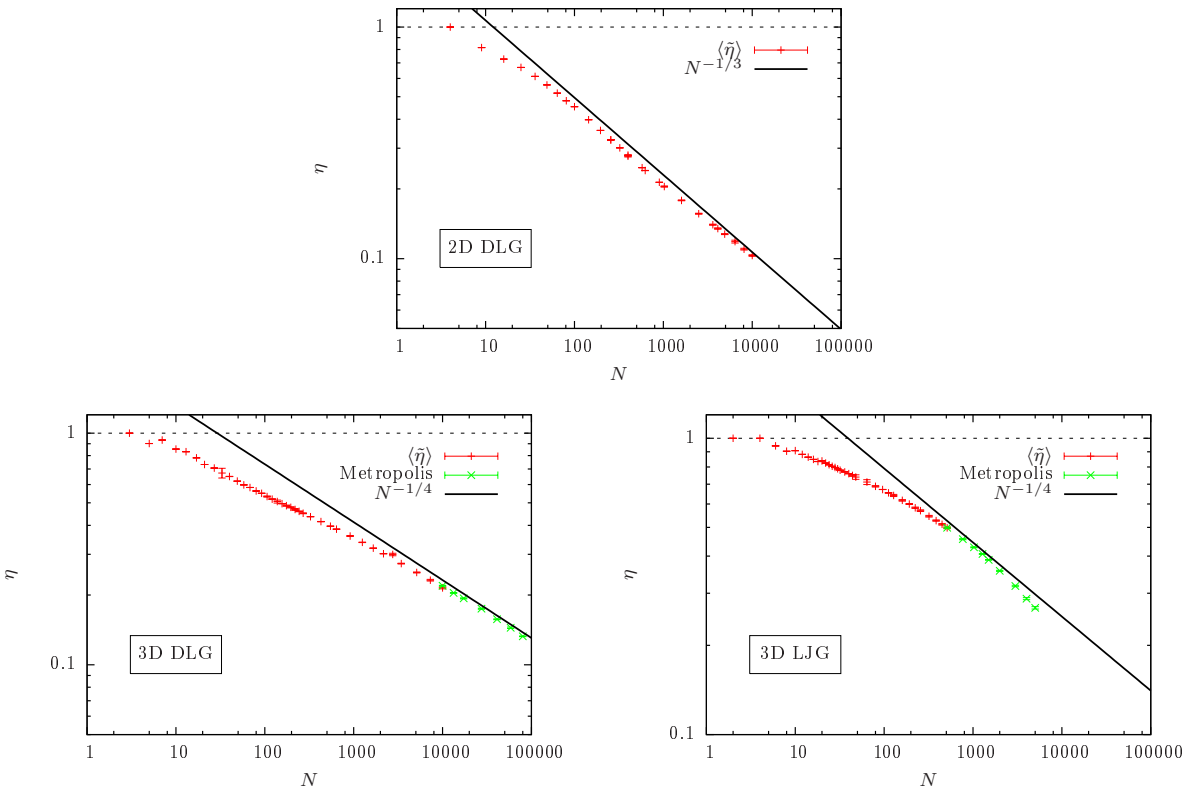

Figure 2.21.: Scaling of the particle fraction η in the largest droplet at the evaporation/condensation transition for the 2D lattice gas (top), 3D lattice gas (bottom, left) and Lennard-Jones gas (bottom, right).

Figure 2.21 shows the result for all three considered models. The fraction $\langle\tilde{\eta}\rangle_{T_c}$ of particles in the transition droplet is obtained as described above from the multicanonical data (red pluses) and from Metropolis simulations (green crosses). The expected scaling behavior is qualitatively recaptured for large system sizes, shown by the (shifted) solid line. For the 2D lattice gas, the expected scaling behavior starts quite early, already for roughly 100 particles. In the 3D cases, however, the expected scaling only starts for quite large system sizes, of the order of 2000 particles for the 3D lattice gas, or may only be anticipated for the 3D Lennard-Jones gas. Including the Metropolis data for the Lennard-Jones gas shows that the leading-order extrapolation of the available data is not precise enough and seems to overestimate the transition temperature for larger systems, consistent with the observations in Sec. 2.3.1.

The dashed line in Fig. 2.21 at $\eta = 1$ corresponds to the case of all particles in the droplet. All three cases show for small systems a majority of particles inside the transition droplet. It needs to be mentioned that for small system sizes, the estimate of the transition droplet via the probability distribution is not very precise, since the distributions are very narrow in $[1, N]$ and sometimes do not even show a clear double peak. However, a reasonable qualitative result is that most constituents of the system end up in the transition droplet. This on the other hand would explain an intermediate scaling regime, where the relevant length scale (the linear extension of the droplet) is now proportional to the linear system size, as argued for flexible polymer aggregation (see Sec. 3.4 and Ref. [42]).

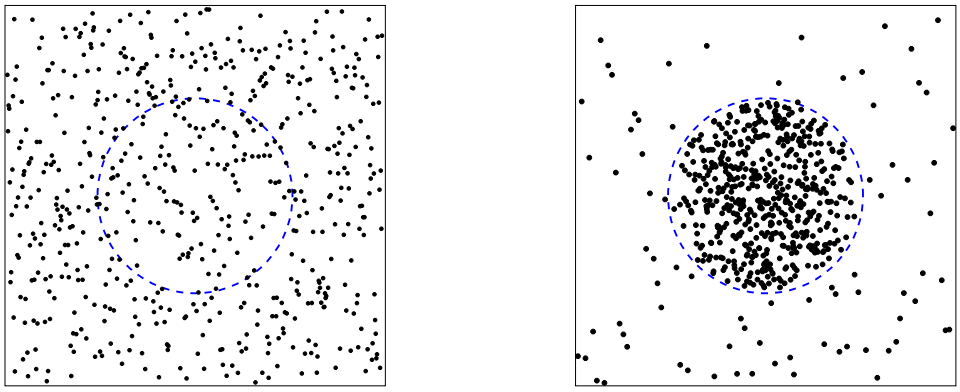

Figure 2.22.: Sketch of a "virtual" subsystem around the idealized transition droplet above (left) and below (right) the condensation/evaporation transition temperature. Within this subsystem, the condensation transition would correspond to a phase transition between homogeneous phases with open boundary conditions.

2.3.4. Relation to non-periodic first-order phase transitions between homogeneous phases

The presented "orthogonal" formulation and subsequent discussion of numerical results show that the predicted scaling behavior for equilibrium droplets at coexistence is valid also in the fixed-ρ approach for large system sizes. In three dimensions it seems that this approach comes even with smaller finite-size corrections, compared to the fixed-T literature and Sec. 2.2. Comparing similarly the two-dimensional lattice gas shows the opposite picture with small finite-size corrections in the fixed-T but large finite-size corrections in the fixed-ρ approach. Thus, it seems that the orthogonal crossing of the phase boundary may lead in certain situations to reduced finite-size corrections and serves as a useful, complementary approach.

The leading-order finite-size scaling corrections may be identified as powers of the linear extension of the transition droplet, which is consistent with the numerical results. Moreover, the involved power-series expansion in the formulation of the finite-size corrections for the temperature and rounding allow to consider effective, empirical next higher-order corrections. The presented data confirm this approach numerically, while deviations from available exact results give a hint that there might be additional corrections of the same order.

In addition, the fixed-ρ approach revealed an intermediate regime, consistent with observations for flexible polymer aggregation (see Sec. 3.4 and Ref. [42]). The presented data allow to argue that this is due to the majority of particles involved in the transition droplet formation. Thus, the droplet size becomes proportional to the system volume and the relevant length scale is the linear system size. This leads to an effective different local scaling behavior. However, including the empirical higher-order term allows to cover the intermediate regime. It may be argued that this will be also the case for the flexible polymer aggregation, if larger system sizes are considered.

An intuitive approach to this may be the competition of finite-size contributions from volume (V^d) and surface (V^{d-1}). This would give rise to a finite-size correction of the order L^{-1}, where L is the relevant length scale of the system [42, 45–47]. It seems natural that the linear extension of the droplet R at coexistence plays this dominant role. The condensation/evaporation

transition clearly connects a homogeneous and a mixed phase in the canonical ensemble. However, within the canonical approach, we may construct a (virtual) subsystem with the volume of the transition droplet [21], see Fig. 2.22. By translational invariance, this subsystem can be always constructed around the largest droplet. Above the transition temperature, this subsystem includes a homogenous gas phase while at and below the condensation transition it is filled by the largest droplet and hence shows a homogenous liquid phase. Thus, this may be interpreted as a grand-canonical transition between homogeneous phases in the virtual system spanned by the volume of the transition droplet. By construction, this virtual volume would have open boundary conditions yielding for the transition between homogeneous phases a finite-size shift of order R^{-1} and a finite-size rounding of order R^{-d}, consistent with rigorous results for non-periodic first-order phase transitions [47]. However, the finite-size scaling of R is already non-trivial, as mentioned above.

2.4. Free-energy barriers at the droplet condensation/evaporation transition

In this section, I briefly discuss the involved free-energy barriers at the droplet condensation/evaporation transition. This shall also serve as an example for the following chapters on polymer aggregation. For an order-of-magnitude comparison, I start with the leading-order analytical results for the fixed-T approach, followed by a phenomenological reasoning for the fixed-ρ scheme with corresponding results for $\rho = 10^{-2}$.

Reconsider the free-energy function Eq. (2.18) at fixed temperature. The free-energy barrier is then related to the difference between the local maximum and the local minimum of $\Phi_\Delta(\lambda) = \lambda^{(d-1)/d} + \Delta(1-\lambda)^2$, at $\lambda_{\max}$ and $\lambda_{\min}$ respectively:

$$\begin{aligned} \Delta F &= \tau_{\mathrm{W}} \left(\frac{\delta N}{\rho_{\mathrm{L}} - \rho_0} \right)^{\frac{d-1}{d}} \left(\Phi_\Delta(\lambda_{\max}) - \Phi_\Delta(\lambda_{\min}) \right) \\ &= \tau_{\mathrm{W}} \left(\frac{\rho - \rho_0}{\rho_{\mathrm{L}} - \rho_0} \right)^{\frac{d-1}{d}} V^{\frac{d-1}{d}} \Delta\Phi, \end{aligned} \tag{2.51}$$

where $\Delta\Phi = \Phi_\Delta(\lambda_{\max}) - \Phi_\Delta(\lambda_{\min})$ is the height of the barrier in Fig. 2.7 and is size-independent. Using the definition of Δ, Eq. (2.19), the free-energy barrier at the condensation transition (at Δ_{c}) is then

$$\Delta F = \tau_{\mathrm{W}}^{\frac{2d}{d+1}} \left(\frac{2\hat{\kappa}\Delta_{\mathrm{c}}}{(\rho_{\mathrm{L}} - \rho_0)^2} \right)^{\frac{d-1}{d+1}} \Delta\Phi \; V^{\frac{d-1}{d+1}}. \tag{2.52}$$

In principle, $\Delta\Phi$ is a constant depending only on the dimension d, and for a fixed temperature all other parameters are constant as well.

At the condensation transition, the value of $\Phi_{\Delta_{\mathrm{c}}}$ at the local minimum is simply Δ_{c}, i.e. $\Phi_{\Delta_{\mathrm{c}}}(\lambda_{\min}) = \Delta_{\mathrm{c}}$ [11]. The local maximum of $\Phi_{\Delta_{\mathrm{c}}}(\lambda)$ may be obtained from the zeroes of the derivative with respect to λ. Setting the derivative of Eq. (2.20) zero yields

$$\frac{2d}{d-1} \Delta_{\mathrm{c}} \lambda^{1/d} (1 - \lambda) = 1. \tag{2.53}$$

The algebraic solution (using for example mathematica or wolframalpha.com) corresponding to the local maximum is

$$\text{2D} \qquad \lambda_{\max} = \frac{2}{3} - \frac{1}{\sqrt{3}} \approx 0.089316 \tag{2.54}$$

$$\text{3D} \qquad \lambda_{\max} = \frac{1}{6}\left(5 - \frac{4}{\left(19-3\sqrt{33}\right)^{1/3}} - \left(19-3\sqrt{33}\right)^{1/3}\right) \approx 0.080357, \tag{2.55}$$

where I already inserted $\Delta_{\mathrm{c}}^{2D} = (1/2)(3/2)^{3/2}$ and $\Delta_{\mathrm{c}}^{3D} = (1/3)2^{4/3}$ from Eq. (2.21). The result for the barrier is

$$\text{2D} \qquad \Delta\Phi \approx 0.1421 \tag{2.56}$$

$$\text{3D} \qquad \Delta\Phi \approx 0.05665, \tag{2.57}$$

see also Ref. [48]. Using the equivalence to the Ising model, the free-energy barrier may be computed for the lattice gas reference cases:

$$\text{2D DLG} \qquad T = 0.375 \qquad \beta\Delta F \approx 0.09592\ L^{2/3} \tag{2.58}$$

$$\text{3D DLG} \qquad T = 0.500 \qquad \beta\Delta F \approx 0.02785\ L^{3/2} \tag{2.59}$$

This is in perfect agreement with the result in Ref. [48], remembering that $\tau_{\mathrm{W}} = \tau_{\mathrm{W}}^{\mathrm{Is}}/4$.

The situation changes, when considering a fixed density. Then, the pre-factor in Eq. (2.52) changes with system size, because it is temperature dependent. However, $\Delta\Phi$ is dimensionless and moreover constant when considering the free-energy barrier. This allows to expand the free-energy difference about the infinite-size condensation transition temperature T_0. Starting from Eq. (2.51), expanding the temperature-dependent part inside the root and recalling $\rho_0(T_0) = \rho$, yields to leading order

$$\begin{aligned}\Delta F &= \left(\tau_{\mathrm{W}}^{\frac{d}{d-1}} \frac{\rho-\rho_0}{\rho_{\mathrm{L}}-\rho_0}\right)^{\frac{d-1}{d}} V^{\frac{d-1}{d}} \Delta\Phi \\ &\approx \left(\frac{d}{dT}\left(\tau_{\mathrm{W}}^{\frac{d}{d-1}} \frac{\rho-\rho_0}{\rho_{\mathrm{L}}-\rho_0}\right)\Big|_{T_0}\right)^{\frac{d-1}{d}} (T-T_0)^{\frac{d-1}{d}} V^{\frac{d-1}{d}} \Delta\Phi.\end{aligned} \tag{2.60}$$

At the condensation transition T_{c}, the free-energy barrier may be then obtained by inserting the leading-order scaling of the transition temperature from Eq. (2.41), $T_{\mathrm{c}} = T_0 + aV^{-1/(d+1)}$. This results in the leading-order scaling of the free-energy barrier

$$\Delta F|_{T_{\mathrm{c}}} \propto V^{(d-1)/(d+1)}. \tag{2.61}$$

On the other hand, analogue to usual free-energy barrier and interface tension discussions (see e.g. Ref. [40] and references therein), one may argue that the barrier should scale as the surface of the relevant interface. As shown in Sec. 2.3, the volume of the transition droplet scales with $V_{\mathrm{D}} \propto N^{d/(d+1)}$. Consequently, the surface of the transition droplet should scale to leading-order with

$$S_{\mathrm{D}} \propto V_{\mathrm{D}}^{(d-1)/d} \propto N^{(d-1)/(d+1)}, \tag{2.62}$$

in accordance with the above results. Thus, I will consider the following scaling ansatz for the free-energy barrier:

$$\beta\Delta F = c_0 + c_1 N^{(d-1)/(d+1)} \tag{2.63}$$

The multicanonical simulation allows to easily measure the free-energy barrier using the reweighted energy probability distribution $P_{\mathrm{eqh}}(E)$ at equal-height temperature T_{eqh}. Since $P(E) \propto \Omega(E)e^{-\beta E}$ and $S(E) = \ln \Omega(E)$ (up to a constant), one may define the microcanonical free energy $F(E) = E - TS(E) = -k_{\mathrm{B}}T \ln P(E)$ up to a constant. The free-energy barrier is then obtained by the difference in free-energy of probable states (maximum in the distribution, minimum in free-energy) and suppressed states (minimum in distribution, local maximum in free-energy), i.e.,

$$\Delta F = k_{\mathrm{B}}T \left(\ln P_{\max} - \ln P_{\min}\right). \tag{2.64}$$

In order to compare to the fixed-T results, the present Ansatz in Eq. (2.63) has to be rewritten in L, namely $\beta\Delta F = c_0 + c_1 \rho^{(d-1)/(d+1)} L^{d(d-1)/(d+1)}$. For $\rho = 10^{-2}$, this leads to the factors $\rho^{1/3} \approx 0.215$ and $\rho^{1/2} = 0.1$ in two and three dimensions, respectively.

Figure 2.23 shows the reweighted energy probability distributions from the multicanonical results of Sec. 2.3 together with the leading-order scaling of the free-energy barrier. The x-axis of the energy probability distribution is shifted to the energy of the gas phase and rescaled in order to account for the leading-order correction of the latent heat $\Delta e = \Delta E/N$ where possible (see e.g. Ref. [20] for the lattice gas case). The free-energy barrier is fitted with the Ansatz in Eq. (2.63) and the lower fit range is indicated by the vertical dashed line.

All three cases show a double-peak energy distribution at the condensation transition with a suppressed region in between - a local minimum. This is a characteristic of first-order phase transitions. This suppression increases with system size, corresponding to an increase in the free-energy barrier. Also, in all cases the scaling of the free-energy barrier is in good agreement with the predicted behavior, shown by a good reduced chi-square normalized with the degrees of freedom ($\chi^2 \approx 1$). A fit to the 2D lattice gas for $L \geq 400$ yields $c_1 = 0.54(1)$ with $\chi^2 \approx 1.9$. Remember, that for this system, the infinite-size transition temperature is $\tilde{T}_0 = 0.39882...$, see Sec 2.3.1. Using the above mentioned conversion, this corresponds to a scaling with linear system size $\beta\Delta F \approx 0.116 L^{2/3}$ which is of the same order as the scaling of the free-energy barrier for $T = 0.375$ in Eq. (2.58). For the 3D lattice gas a fit to $L \geq 40$ yields $c_1 = 0.566(2)$ with $\chi^2 \approx 2.3$. Again, this may be rewritten as $\beta\Delta F \approx 0.0567 L^{3/2}$, which is quite larger than the fixed-T scaling in Eq. (2.59) for $T = 0.5$. As a final result, a fit to the 3D Lennard-Jones gas for $N \geq 22$ yields $c_1 = 1.016(2)$ with $\chi^2 \approx 0.7$. This is even larger than for the 3D lattice gas, while the achievable system sizes are a lot smaller.

From the present results, it may be argued that the finite-size corrections to the first-order transition decrease with increasing suppression of intermediate states and consequently with an increasing free-energy barrier. The free-energy barrier is a measure of the first-order signature for the finite system. A larger slope in the scaling of the free-energy barrier thus means that the system shows the leading-order behavior already for smaller system sizes, approaching the thermodynamic limit “faster”. This would explain why the 3D lattice gas shows smaller finite-size corrections for fixed-ρ (Sec. 2.3) than in the corresponding fixed-T case (Sec. 2.2). The relatively larger barrier for the Lennard-Jones gas also explains why – despite much smaller system sizes – the finite-size scaling analysis in Sec. 2.3 shows the qualitative leading-order scaling behavior already for comparably smaller system sizes as in the 3D lattice gas.

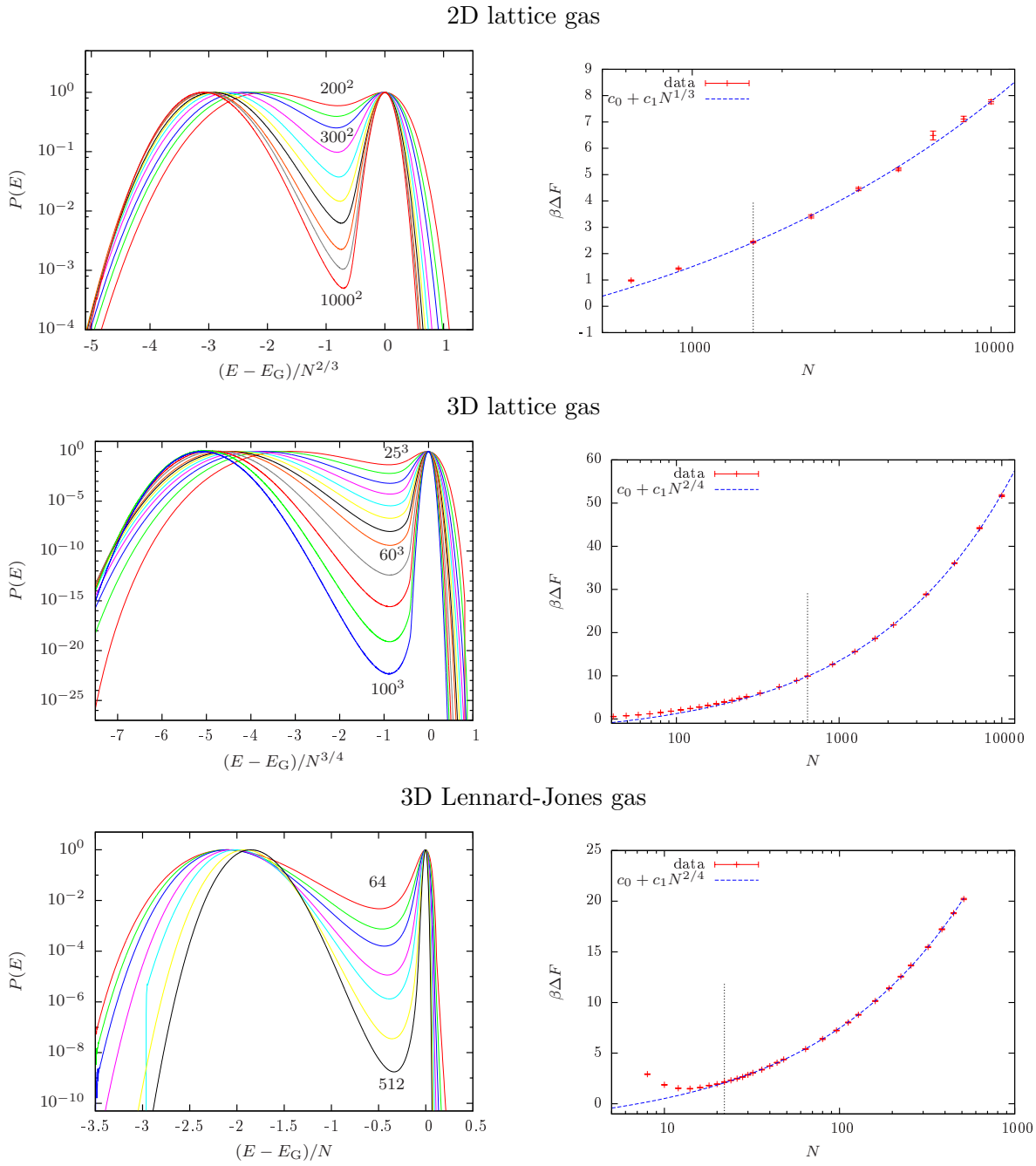

Figure 2.23.: Reweighted energy probability distribution (left) at equal height for several system sizes in steps of $\Delta L = 50, 100$ (2D lattice gas), $\Delta L = 5, 10$ (3D lattice gas) and $N = 2^{(n/2)}$ (Lennard-Jones gas). Scaling of the free-energy barrier (right) as defined in the text. The vertical dashed line indicates the lower fit range.

2.5. Finite-size scaling in the microcanonical ensemble

The microcanonical analysis [49, 50] has been shown to be a powerful complementary tool to the standard canonical analysis, see for example Refs. [51–55]. In the thermodynamic limit, the canonical and microcanonical ensemble are expected to coincide. However, for finite systems both ensembles provide different perspectives on the same problem. Certain aspects will be more clear in one or the other ensemble, depending on the question at hand. In this section, I will present a short introduction to the microcanonical ensemble and the corresponding definitions relevant for a microcanonical analysis. Using particle condensation as an intuitive example for a first-order phase transition, I will present characteristic signatures and systematic finite-size scaling in the microcanonical ensemble. This shall again serve as a preparation for the less-well understood case of polymer aggregation.

In the microcanonical (NVE) ensemble, the energy E is fixed and the temperature T becomes an observable. In general, the NVE ensemble is defined for a fixed total energy E_t, i.e., the sum of kinetic and potential energy. Of course, one may perform corresponding microcanonical Monte Carlo simulations [56–59]. However, one may also consider the microcanonical ensemble formulated for a fixed potential energy, if the (isotropic) kinetic part may be formally integrated out, see also Sec. 5.1. Then, the observables become functions of the potential (conformational) energy $E = E_p$. Estimates of the corresponding (conformational) microcanonical entropy and inverse temperature may be obtained directly from canonical energy distributions [50]. The connecting function is the density of states. On the other hand, the multicanonical method and other generalized ensemble simulations yield a direct estimate of the density of states as natural output (see Sec. 5.4). Thus, an analysis in the microcanonical ensemble comes with little additional effort, if the estimate of the density of states is sufficiently precise. Considering a multicanonical simulation as defined in Sec. 5.4, the weight function $W(E)$ may be directly related to the density of states $\Omega(E)$ as $\Omega(E) \approx W^{-1}(E)$. Taking into account the final production run with histogram $H(E)$, one may increase the accuracy of the density of states with increasing the number of measurements, where $\overline{\Omega}(E) = H(E)W^{-1}(E)$ (up to a multiplicative constant). For technical details see Sec. 5.4.1. This will serve as the starting point for the more general microcanonical analysis based on the density of states.

The microcanonical entropy is defined as $S(E) = k_\mathrm{B} \ln \Omega(E)$, where $k_\mathrm{B} = 1$ in the following. It is a fundamental property of each physical system and a strictly concave function in the thermodynamic limit. For finite systems showing transitions with phase separation it may show convex regions, or *convex intruders* [49]. Figure 2.24 (left) shows the microcanonical entropy with convex intruder using the example of a finite ($N = 32$) particle Lennard-Jones gas, where for larger system sizes the convex intruder becomes less prominent. This is an intuitive example for a transition with phase separation, where a homogeneous (gas) phase and an inhomogeneous (condensate plus gas) phase coexist at the transition temperature. The vertical dashed lines show the maxima of the canonical energy distribution at the condensation transition, here defined with two peaks of equal height determined in Sec. 2.4, with the corresponding energy of the condensate phase E_cond and of the gas phase E_gas. The concave hull $\mathcal{H}_S(E) = S(E_\mathrm{cond}) + E/T_\mathrm{cond}$ is the tangent connecting $S(E_\mathrm{cond})$ and $S(E_\mathrm{gas})$ and corresponds to the Gibbs construction [51]. In fact, this condensation transition temperature corresponds to the temperature T_eqh for the canonical equal-height energy distribution, because then $P(E_\mathrm{cond}) = P(E_\mathrm{gas})$ [50]. In general, the derivative of $S(E)$ at fixed volume V and

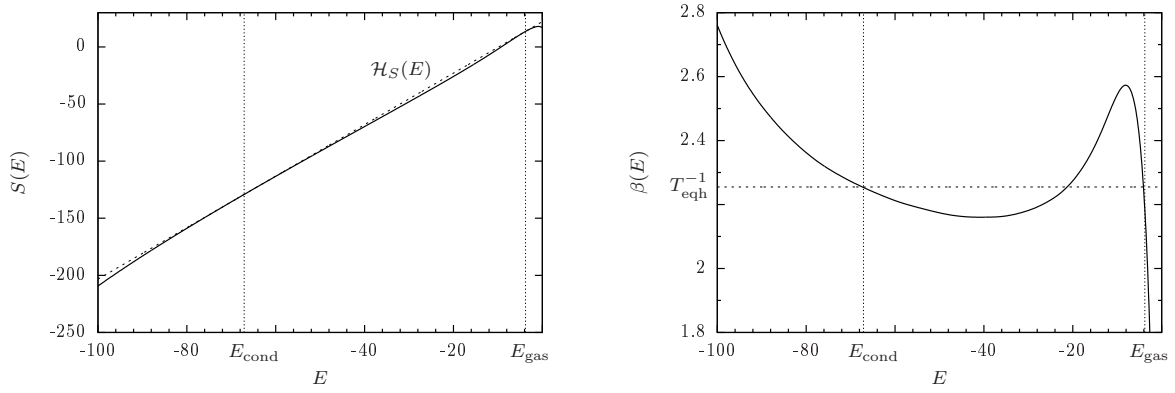

Figure 2.24.: Example of a microcanonical entropy $S(E)$ (left) and inverse temperature $\beta(E)$ (right) for a finite $N = 32$ particle Lennard-Jones ($\rho = 10^{-2}$) system with a phase-separation transition. $S(E)$ shows the characteristic convex intruder and $\beta(E)$ the corresponding backbending effect. The dashed lines show the results from the equal-height histogram in Sec. 2.4.

particle number N is the microcanonical inverse temperature

$$\beta(E) = \left(\frac{\partial S(E)}{\partial E}\right)_{N,V}. \tag{2.65}$$

As long as the microcanonical entropy is a concave function, this inverse temperature is a monotonically decreasing function in E (however it may become negative if defined in terms of the potential energy). The convex intruder, on the other hand, causes a non-monotonic inverse temperature curve with a locally increasing region, also referred to as the *backbending effect* in the microcanonical inverse temperature. This is shown for the considered example Fig. 2.24 (right). The inverse condensation transition temperature from the convex hull or the canonical equal-height energy distribution intersects the microcanonical inverse temperature curve at two characteristic points: the condensate energy and the gas energy. It thus allows to estimate the latent heat $\Delta Q = E_{\mathrm{gas}} - E_{\mathrm{cond}}$ which is the energetic width of the transition region. By construction, both enclosed areas have the same size.

Following Refs. [49, 55], the occurring (structural) transitions in finite systems may be classified by considering the derivative of the microcanonical inverse temperature:

$$\gamma(E) = \left(\frac{\partial \beta(E)}{\partial E}\right)_{N,V} = \left(\frac{\partial^2 S(E)}{\partial E^2}\right)_{N,V} \tag{2.66}$$

Transitions are defined *first order* for a local maximum $\gamma(E_{\mathrm{tr}}) > 0$, which corresponds to a positive slope of $\beta(E)$ at the corresponding inflection point, necessary for the backbending effect. Then, there exists a latent heat and physical phase coexistence. Analogously, transitions are defined *second order* for a local maximum with $\gamma(E_{\mathrm{tr}}) < 0$, i.e., a negative slope of $\beta(E)$ with no latent heat and no coexistence. In a finite-size scaling towards the thermodynamic limit, both local maxima should approach zero: for a first-order transition from above and for a second-order transition from below. In this case, the notion of a phase transition in the usual sense is recovered.

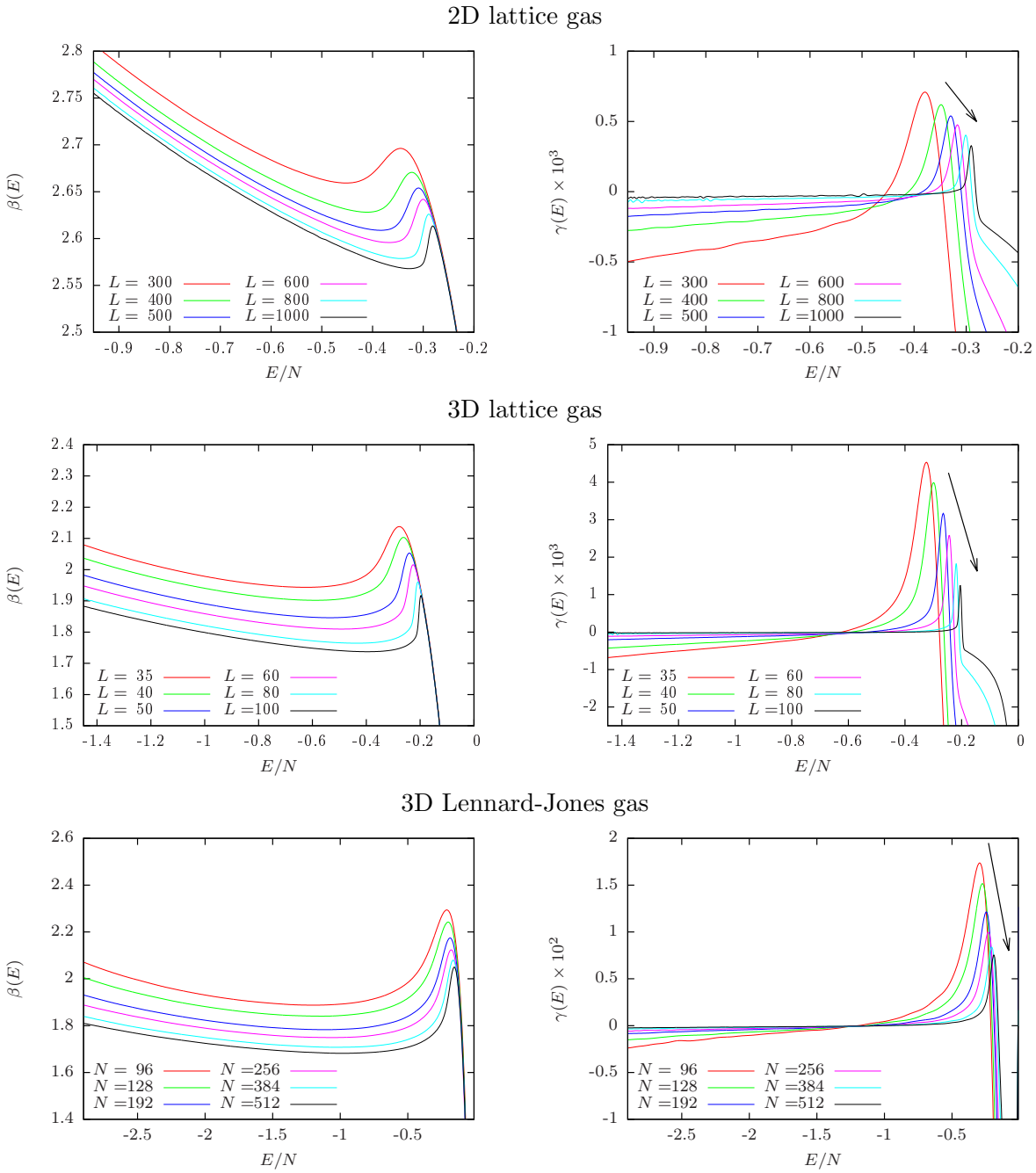

Figure 2.25.: Microcanonical analysis for all considered particle gas models ($\rho = 10^{-2}$) using the first (left) and second (right) derivative of the microcanonical entropy. For a first-order phase transition, the positive peak in $\gamma(E)$ approaches zero from above in the thermodynamic limit.

Examples for first- and second-order structural transition signatures in finite systems are shown in Ref. [55]. In order to illustrate the finite-size behavior for a first-order phase transition, I reconsider the particle condensation examples at $\rho = 10^{-2}$ from the previous section, shown in Fig. 2.25. The microcanonical inverse temperature $\beta(E)$ shows the characteristic backbending effect, resulting in a positive slope at the inflection point and consequently a positive peak in $\gamma(E)$. In order to display various system sizes with up to 10 000 lattice particles and 512 Lennard-Jones particles, the average energy per particle is plotted. This may lead to the (false) perception that the local slope in $\beta(E)$ increases with system size, while in fact the derivative with respect to the energy decreases with system size, shown by the black arrows in the corresponding figures of $\gamma(E)$. Thus, the maxima seem to fulfill $\gamma_N(E_{\rm tr}) \to +0$ from above for $N \to \infty$ and the notion of a first-order phase transition for particle condensation remains valid in the thermodynamic limit.

2.6. Combining temperature and density

So far in the discussion, I considered only a single density $\rho = 10^{-2}$ with the assumption that the arguments hold for any density small enough to show a clear condensation/evaporation transition. The density will, however, strongly influence the finite-size scaling limit. This may be demonstrated using the equivalence of the density ρ to the spontaneous magnetization m_0 in the Ising model, see also Sec. 2.1.1 and Fig. 2.13. For an illustration consider the low-temperature series expansion of the three-dimensional Ising model. For low temperatures (and small densities) the spontaneous magnetization is given to first order as the series expansion $m_0 = 1 - 2e^{-12\beta^{\rm Is}} + ...$, where $\beta = 4\beta^{\rm Is}$. This may be read as a relation between the infinite-volume transition temperature and the fixed system density $\rho = \frac{1}{2}(m_0(T_0) - 1)$. For $T_0 = 1/\beta_0 \to 0$ or $\rho \to 0$, this yields $\rho = e^{-3\beta_0}\,(1 + ...)$, and to first order for the inverse transition temperature of the infinite system

$$\beta_0 \simeq -\frac{1}{3} \ln \rho. \tag{2.67}$$

This result may be similarly deduced from microcanonical arguments in continuous systems (see Ref. [42]). To this end, recall the discussion of the microcanonical inverse temperature in Sec. 2.5. The inverse (finite) condensation temperature may be estimated as the slope of the hull that is tangent to the microcanonical entropy at both the condensate and the gas phase:

$$\beta_{\rm cond} = \frac{S(E_{\rm gas}) - S(E_{\rm cond})}{\Delta E} \tag{2.68}$$

Now, consider a variation in density ρ shown on the example of $N = 30$ Lennard-Jones particles in Fig. 2.26. The microcanonical entropy (left), estimated from the multicanonical simulation as in Sec. 2.5, is almost unaffected in the droplet regime at low energies but strongly density-dependent in the gas regime at high energies. Of course, this is a very small system explaining the nucleation effects at high energies, which become more pronounced for smaller densities. Also, this small system includes a large fraction of particles in the largest droplet, even increasing with decreasing density, see Fig. 2.26 (right). The finite-size scaling of the droplet size should additionally lead to a density dependence of the condensate regime with increasing system sizes and decreasing density. However, the entropy change in the gas phase should clearly dominate any entropy change in the droplet phase. The microcanonical entropy is proportional to the logarithm of the number of states with a given energy, $S(E) = k_{\rm B} \ln \Omega(E)$

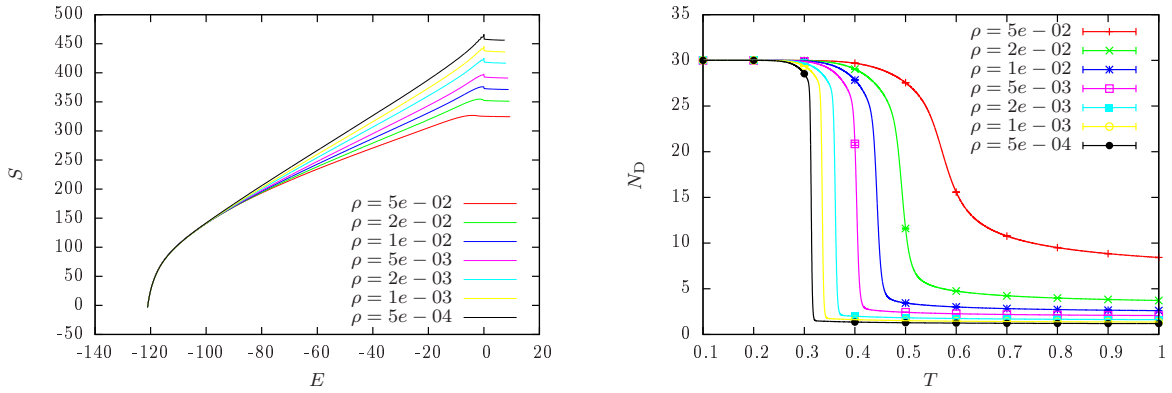

Figure 2.26.: Example of the microcanonical entropy up to an additive constant (left) and the number of particles N_D in the largest droplet (right) for the $N = 30$ 3D Lennard-Jones gas.

where $k_B = 1$ here. Assuming ideal-gas like behavior in the gas regime, the number of states may be approximated as V^N which leads to

$$S(E_{gas}) \sim \ln\left[V^N\right] \propto -N \ln \rho, \tag{2.69}$$

for a constant number of particles with $\rho = N/V$. This will be much larger than $S(E_{cond})$, where a fraction of particles is in the droplet and the remaining fraction still in the gas phase. For a fixed number of particles ΔE may be considered constant and the inverse condensation temperature thus becomes a function of the density

$$\beta_{cond} \sim \frac{S(E_{gas})}{\Delta E} \sim -c_1 \ln \rho + c_2, \tag{2.70}$$

see also similar arguments for lattice proteins [60] and polymer adsorption [54]. These arguments apply also for finite systems, where $\{c_1, c_2\}$ should of course be system-size dependent.

At this point, I want to reconsider the finite-size scaling analysis from Sec. 2.3.1 for several fixed densities. Figure 2.27 (left) shows the finite-size transition temperatures for the 3D Lennard-Jones gas with the corresponding higher-order fits. Due to the model with cutoff, however, the explicit values of T_0 deviate from real examples [43]. In general, the Lennard-Jones gas may be reasonably applied to non-polar gases, like Argon. In this case, molecular dynamics simulations date back to the 60s [61]. Corresponding parameters are $\sigma \approx 3.4\mathring{A}$ and $\epsilon/k_B \approx 120K$ and serve well for an order-of-magnitude comparison. Then, $T_{real} = T \times 120K$ and $\rho_{real} \approx \rho \times 1.226 \times 10^3 g/l$, see also Ref. [21]. Thus, the Argon boiling point $T_{boiling} \approx 87.3K$ [62] with a gas density $\rho_{boiling} \approx 5.772 g/l$ at atmospheric pressure. This corresponds to $T_0' \approx 0.728$ and $\rho' \approx 0.005$ in the current dimensionless units. The finite-size scaling extrapolation with higher-order corrections for the corresponding density $\rho = 5 \times 10^{-3}$ yields $T_0 = 0.645(3)$ which differs from the experimental result by $\sim 10\%$, as expected for the truncated model. The extrapolated infinite-size condensation transitions T_0 are shown as a function of ρ in Fig. 2.27 (right). In fact, this is well described by the inverse of Eq. (2.70), namely $T_0(\rho) = 1/(c_2 - c_1 \ln \rho)$ with a reduced $\chi^2 \approx 1.2$ for $\rho < 0.03$. The corresponding fit result is $c_1 \approx 0.173$ and $c_2 \approx 0.634$

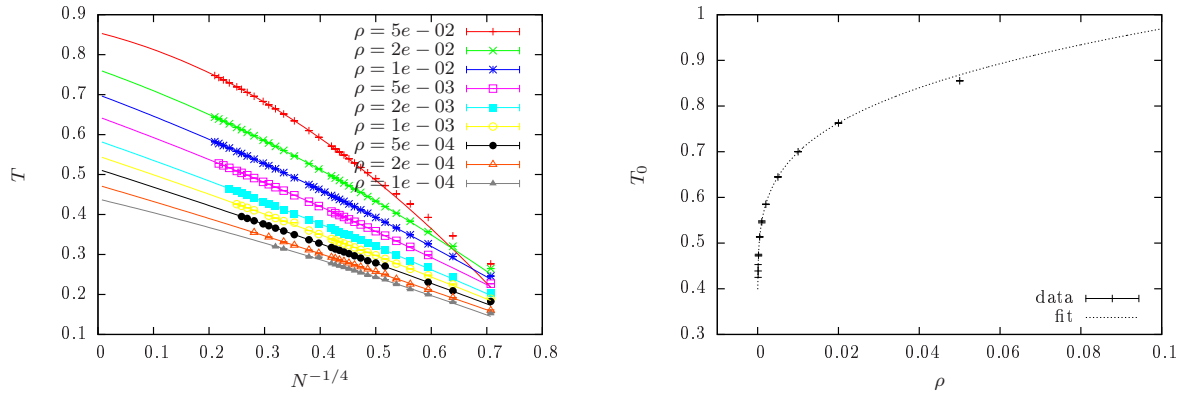

Figure 2.27.: Finite-size scaling of the condensation transition temperature (left) for the 3D Lennard-Jones particle gas at several densities $\rho = N/V$. The extrapolation to the infinite-sized system is shown on the right with a fit according to the inverse of Eq. (2.70).

Focusing on the finite-size corrections, Fig. 2.28 (left) shows the transition temperature rescaled with the density dependence of T_0. With decreasing density, the rescaled temperature curves seem to collapse and follow a joint straight line. However, for a wide range of densities, the finite-size scaling corrections are density dependent. Considering the ansatz $T_c = T_0 + aN^{-1/4} + bN^{-2/4}$ from Sec. 2.3, the finite-size corrections may be expressed in terms of the infinite-system transition temperature, namely a/T_0 and b/T_0 shown in Fig. 2.28 (right). With decreasing density the leading-order correction amplitude seems to saturate to a finite value and the effective higher-order correction amplitude seems to vanish. This observation justifies to consider only an effective leading-order correction for sufficiently small densities. The leading-order correction amplitude is shown in the figure as a_{leading}/T_0 and is always a little smaller as when including higher-order corrections. A rough estimate for the effective leading-order correction for sufficiently small densities is $a_{\text{leading}}/T_0 \approx 0.94$

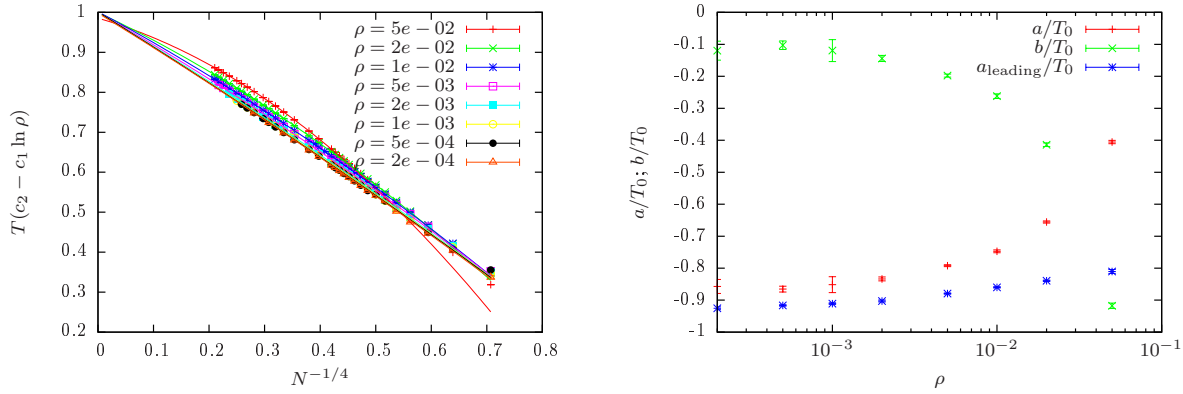

Figure 2.28.: Finite-size scaling of the rescaled condensation transition temperature (left) together with the finite-size correction terms a/T_0, b/T_0 and a_{leading}/T_0 for the leading-order fit as a function of density (right).

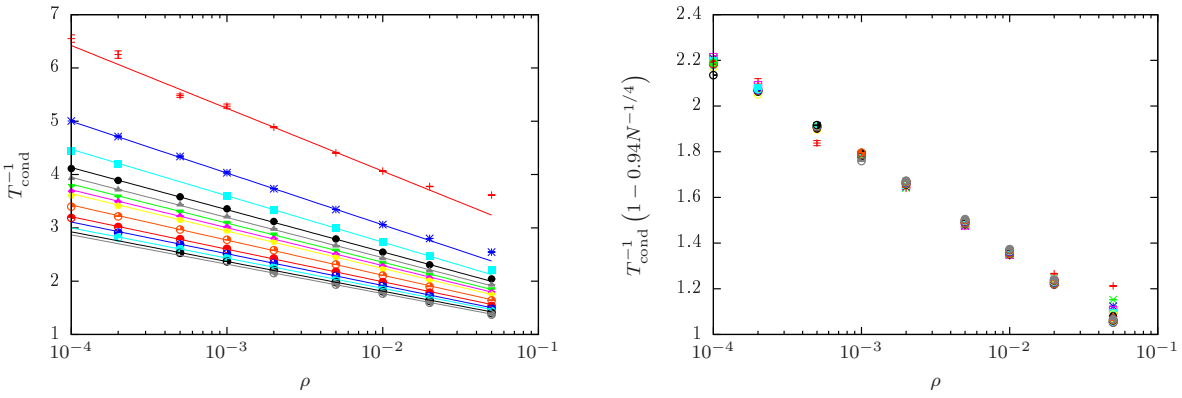

Figure 2.29.: Density dependence of the finite condensation temperature of the 3D Lennard-Jones gas (left) for $N = 4$ (upper curve) until $N = 320$ (lower curve). The same data with rescaled temperature shows a data collapse (right), considering an effective leading-order finite-size scaling for sufficiently small densities.

The logarithmic dependence of the (finite) inverse condensation transition is shown in Fig. 2.29 for a fixed particle number and varying system volume. Starting with $N = 4$ for the upper curve, the system size gradually increases up to $N = 320$ for the lower curve in the left panel. All system sizes show the predicted density dependence, shown by the fits to the data. However, the fit constants show to be slightly system dependent, especially for small number of particles. Considering in addition the effective leading-order finite-size scaling of the transition temperature yields a data collapse for the density dependence for a broad range of density and particle number in Fig. 2.29 (right). Of course, there are deviations for both large densities and small particle number, but the overall behavior seems to be nicely explained by considering both leading-order entropic and energetic arguments.

3. Aggregation transition in dilute systems with flexible homopolymers

Homogeneous aggregation of dilute polymers describes the transition between separated, individual polymers and a macroscopic aggregate. For an aggregate in a finite system there may or may not remain unattached polymers, depending on the specific system setup (system size, density, solvent, ...). In the limit of increasing polymer number, this is expected to be analogue to particle condensation, where the aggregation transition of *dilute* polymer systems separates a homogeneous phase of "isolated" polymers from a mixed phase with a single macroscopic aggregate and surrounding unattached polymers. Moreover, one may explore the similarity between simple liquids, polymer solutions and single polymers in appropriate limits, see e.g. Refs. [2, 3, 63]. For the limit of increasing polymer number at fixed polymer length, this will be discussed in detail in Sec. 3.4, building on the results for particle condensation and discussing a qualitative crossover and arising similarities. However, most system sizes remain within a *mesoscopic* regime, which may be related to the intermediate scaling regime in particle condensation.

In addition, it is of fundamental interest to understand the underlying mechanisms of polymer aggregation, which calls for detailed studies of small system sizes including merely a few polymers [7]. In this context, it was shown that aggregation of peptides (heteropolymers) and homopolymers may be considered as a phase-separation process [51–53]. Similarly, for 2-4 rigid-bond (bead-stick) polymers it was shown that aggregation is accompanied by size-dependent hierarchical sub-transitions [53, 64]. For homopolymer systems, the ratio of intra- and inter-polymer interaction strength is often considered to be 1. Deviations from this choice lead to a shift in aggregation temperature and may pronounce or suppress the single-polymer collapsed "phases" (see Sec. 3.2). The interplay between the general case of polymer collapse and aggregation at equal interaction strength shows that aggregation dominates (and may even revert) the continuous collapse transition [42, 65]. Also, using entropic and energetic arguments, it is possible to describe the aggregation temperature in terms of the density and polymer number (see also Sec. 3.5). For a specific sequence of lattice heteropolymers, the interplay between folding and structural assembly was studied in Ref. [60], observing a similar relation between the transition temperature and density. This supports that the underlying mechanism is quite generic.

This chapter is organized as follows: in Sec. 3.1, I will introduce the considered polymer models – the interacting self-avoiding walk and a bead-spring polymer – and compare with exact results and Metropolis simulations. In Sec. 3.2, I discuss the effect of intra- and inter-polymer interaction using exact results. For the remaining part I am restricting myself to the case of equal interaction strengths. Possible finite-size scaling limits are discussed in Sec. 3.3. Here, I revisit the collapse transition of a single polymer in Sec. 3.3.1, demonstrating the equivalence to polymer models in the literature followed by an example of the $N \to \infty$ limit for polymer aggregation in Sec. 3.3.2. The main finite-size scaling results are analyzed in the $M \to \infty$ limit in Sec. 3.4, including the canonical aggregation temperature, transition

rounding and size of the transition aggregate; a microcanonical analysis and the discussion of a qualitative crossover from particle condensation to (short) polymer aggregation. In the last part, Sec. 3.5, I discuss the dependence of polymer aggregation on density, identifying the crossover from semi-dilute to dilute solutions and a relation between the transition temperature and polymer density. Considering in addition finite-size-scaling results leads to a data collapse.

3.1. Coarse-grained polymer models

In this chapter, I mainly consider two simple coarse-grained homopolymer models. As a natural generalization from the study of particle condensation, these include the discrete interacting self-avoiding walk (iSAW) and the bead-spring polymer model. Both models share a self- and mutual-avoidance plus short-range attraction. The first model is a lattice realization with nearest-neighbor interaction, similar to the lattice gas example. The second case is one of the popular generic polymer models with non-bonded Lennard-Jones (LJ) interaction and *harmonic* bonds, thus for certain aspects comparable to the Lennard-Jones gas. For a sketch of the models and interaction potentials see Fig. 3.1.

Again, the canonical ensemble is the natural choice, i.e., a fixed temperature T, polymer number M, and system volume $V = L^3$ with linear system size L. However, polymers are now extended objects, here linear chains with N connected monomers. Note the change of notation from particle condensation, where N_{gas} refers to the number of particles. While this is a little inconvenient, it is chosen in order to be consistent with the literature on both particle condensation and polymer aggregation. I consider uniform polymer systems, i.e., M polymers of the same length N, instead of polydisperse systems. This is because I am interested in general features of aggregation and the emerging parameter space will show to be sufficiently large already for uniform systems.

In general, the observables that I consider include as usual the total (potential) energy E and its derivative, the specific heat per monomer

$$C_V = k_B\beta^2 \left(\langle E^2\rangle - \langle E\rangle^2\right)/NM, \tag{3.1}$$

with the inverse temperature $\beta = (k_{\text{B}}T)^{-1}$. The Boltzmann constant is set to unity in general $k_{\text{B}} = 1$, which sets the temperature scale. In order to bridge the gap to particle condensation, I consider the number of particles in the largest aggregate (droplet) N_{D}. Following Refs. [52, 53], I consider also a "phase" separation parameter defined as radius of gyration of the center of masses $\boldsymbol{r}^i_{\text{cm}}$ of individual polymers,

$$\Gamma^2 = \frac{1}{2M^2}\sum_{i,j}\left(\boldsymbol{r}^i_{\text{cm}} - \boldsymbol{r}^j_{\text{cm}}\right)^2. \tag{3.2}$$

3.1.1. Lattice polymer

The lattice polymer model that I consider is sometimes also referred to as the interacting self-avoiding walk (iSAW). Each monomer is described by an occupied site on a lattice (here the 3D cubic lattice), where each site may be occupied only once. This accounts for the hard-core repulsion, or self- and mutual-avoidance. For each polymer, the monomers are connected by stiff bonds along the edges of the lattice, such that the bond length is the lattice spacing a.

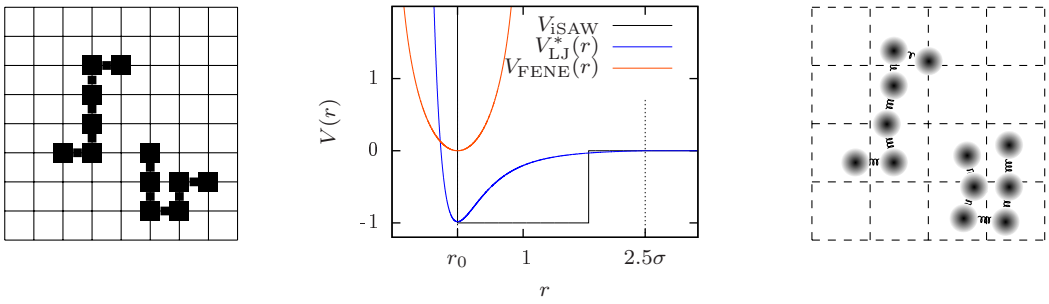

Figure 3.1.: Sketch of two complementary polymer models: the discrete lattice polymer model (left) and the continuous bead-spring polymer model (right). The corresponding non-bonded (iSAW,LJ) and bonded (FENE) interaction potentials are plotted in the center.

A short-range attraction is acting only between non-bonded nearest-neighbors. For $a = 1$, the corresponding interaction potential $V_{\mathrm{iSAW}}(r)$ assumes the values $V_{\mathrm{iSAW}}(0) = \infty$, $V_{\mathrm{iSAW}}(1) = -\epsilon$ and zero else. This is identical to V_{DLG} in the lattice gas case. Of course, ϵ may depend on the specific sort of interaction, for example $\epsilon_{\mathrm{intra}}$ and $\epsilon_{\mathrm{inter}}$ for intra- and inter-polymer interaction, respectively. Expressed in terms of number of intra-polymer contacts n_{intra} and inter-polymer contacts n_{inter} the Hamiltonian then reads

$$\mathcal{H} = -\left(\epsilon_{\mathrm{intra}} n_{\mathrm{intra}} + \epsilon_{\mathrm{inter}} n_{\mathrm{inter}}\right). \tag{3.3}$$

For most purposes I consider $\epsilon_{\mathrm{intra}} = \epsilon_{\mathrm{inter}}$. However, Sec. 3.2 includes a qualitative discussion of the effect of the ratio of interaction strengths on the structural transitions of a very small system using exact enumerations. For a single iSAW, a thorough study of the collapse and freezing transitions may be found in Ref. [66]. The majority of my results are obtained from parallel multicanonical simulations (see Sec. 5.4). The Monte Carlo updates include pivot, translation, slithering snake, corner flip and double-bridge moves (for details see Sec. 5.7).

A major feature of lattice models is that one may compare numerical results for small system sizes to exact results, where the density of states is fully enumerated. The outline of the exact enumeration for a single iSAW and 2 iSAWs is presented in Sec. 5.2 together with the reweighting scheme. This allows to compare, in general, most observables of interest. For a quantitative order-of-magnitude comparison of the applied update moves, I performed parallel multicanonical simulations (with 4 cores) of a 2-polymer system with $N = 8$ monomers each (cubic box of linear size $L = 31$). The iterative part uses the full set of updates (fixed sequence), followed by several production runs for different subsets of updates. Considered are the specific heat and the temperature derivative of the phase separation parameter (not normalized in the same way). The canonical estimates $\overline{O}$ and expectation values $\langle O \rangle$ are a function of the temperature, in this case for discrete values in the range $T \in [0.1, 3]$ with $\Delta T = 0.1$. In order to obtain a scalar representative for a comparison of the full temperature range, I consider the following (temperature) sums: the sum of deviations of the estimates from the expectation values $\sum_T |\langle O \rangle(T) - \overline{O}(T)|$ and the sum of (binning) errors of the estimates $\sum_T \epsilon_B\left(\overline{O}(T)\right)$. The exemplary results for single simulations in Table 3.1 show that in almost all cases the total deviation from the exact result is smaller than the statistical error of the

Table 3.1.: Comparison of exemplary multicanonical simulations with exact enumeration using different sets of updates for 2 iSAWs ($N = 8$). Examples for $N = 9$ are shown in Fig. 3.2 and Fig. 3.3. The table shows the temperature-sums of the absolute deviation and statistical error. Considered updates are the pivot (P), translation (T), slithering snake (S), corner flip (F) and the double-bridge (B) move. All cases show that the total deviation and the total error are of the same order of magnitude.

updates	$\sum_T \lvert\langle C_V\rangle - \overline{C}_V\rvert$	$\sum_T \epsilon_B(\overline{C}_V)$	$\sum_T \lvert\langle \frac{d}{dT}\Gamma^2\rangle - \overline{\frac{d}{dT}\Gamma^2}\rvert$	$\sum_T \epsilon_B(\overline{\frac{d}{dT}\Gamma^2})$
P	0.52	0.57	4.96	5.30
P,T	1.07	0.53	5.45	3.63
P,S	0.44	0.75	3.05	5.90
P,F	0.60	0.58	3.16	5.87
P,B	0.33	0.59	3.43	6.06
P,T,S,F,B	0.26	0.64	3.31	5.59

estimates. All update sets include the pivot move because it ensures ergodicity [67]. Only the case of combining the pivot with the polymer translation update, shows a stronger total deviation than total error. Inspecting the full temperature range one can see that this is due to bad sampling in the aggregate, while the high-temperature regime shows smaller deviations. This is explained by the enhanced sampling of translational entropy, a benefit of the translation update. The overall picture, however, indicates that all considered sets of update moves sample the phase space ergodically and do not violate detailed balance.

In order to provide a feeling for the involved (finite) transitions and estimated errors, I continue with a qualitative comparison for 2 polymers of length $N = 9$ ($L = 35$). The temperature estimates are obtained by time-series reweighting of 2.56 million measurements with 50 sweeps per measurement from parallel multicanonical simulations (with 4 cores). A sweep is in general

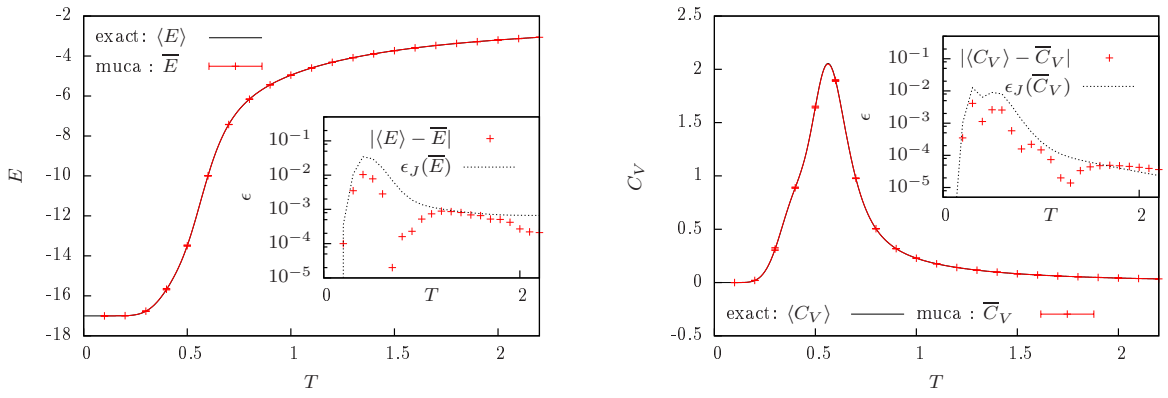

Figure 3.2.: Comparison of a parallel multicanonical simulation with exact enumeration for a system of 2 lattice polymers of length $N = 9$. The reweighted energy (left) and its derivative the specific heat (right) show good agreement with the exact data. The insets show the absolute deviation from the exact expectation values compared to the calculated jackknife error.

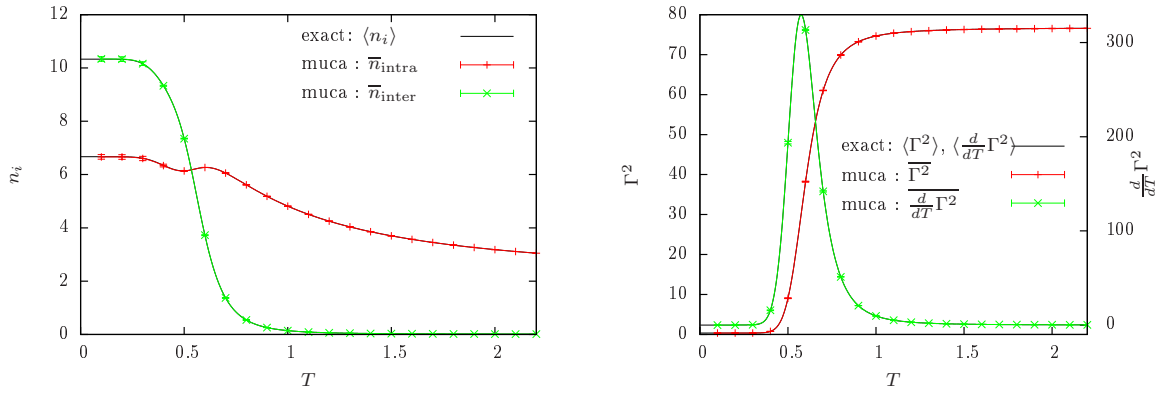

Figure 3.3.: Same as Fig. 3.2 but for the intra- and inter-polymer contacts (left) and the phase separation parameter Γ^2 with its thermal derivative (right). The insets are omitted for clarity but show the same qualitative picture as for the energy.

defined as $N \times M$ updates, randomly chosen from a set of update moves. Figure 3.2 shows the estimated energy and specific heat compared to the exact expectation value. As expected, the energy decreases when lowering the temperature, with a maximal slope at the *aggregation* temperature indicated by a peak in the specific heat. The insets each show the absolute deviations of observables (crosses) compared with the estimated error from jackknife error analysis (dashed line), which was considered above in the quantitative comparison. It is important to keep in mind that the multicanonical simulation yields correlated canonical estimates, because all estimates are computed from the same set of data. Thus, one may not expect a scattering of data points around the estimated error as one would expect for Metropolis simulations. However, it may be seen that for most parts the absolute deviation from the exact solution is small and below the estimated jackknife error. Also, the error and deviation both increase in the vicinity of the transition.

Moving the focus to a qualitative understanding of the typical temperature-dependence of observables, Fig. 3.3 shows the number of intra- and inter-polymer contacts (left) as well as the previously mentioned phase separation parameter (right). It can be nicely observed that with decreasing temperature, the number of contacts increase, corresponding to a decrease in energy. For larger temperatures, the increase in intra-polymer contacts dominates, signaling the onset of a single-polymer collapse. However, in the vicinity of the aggregation transition, the inter-polymer contacts quickly increase and at the same time the intra-polymer contacts slightly decrease before increasing again in order to form the size-specific low-energy conformation. This will be also subject of discussion in Sec. 3.5. The phase separation parameter shows to be a good order-parameter for the aggregation transition. For small temperatures, where the polymers have maximal number of inter-polymer contacts, the phase separation parameter is small. This indicates the aggregated “phase” or regime. With increasing temperature, the polymers separate and the phase separation parameter becomes large until saturation. This is due to a large average distance of independent polymers at high temperature, where entropy dominates energy. The temperature derivative of the phase separation parameter shows a strong peak at the aggregation temperature.

3.1.2. Bead-spring polymer

The bead-spring polymer model is a commonly applied coarse-grained homopolymer model in continuous space. While the involved potentials may vary in detail and parameterization, the essential features are self- and mutual- avoiding monomers (beads) with short range attraction, which are bonded by a (locally) harmonic potential along a single polymer. For the here presented studies, I considered the finitely extensible nonlinear elastic (FENE) potential for *bonded* monomers with a local effective harmonic behavior but with a maximal and minimal bond length:

$$V_{\mathrm{FENE}}(r) = -\frac{K}{2} R^2 \ln\left(1 - [(r - r_0)/R]^2\right), \tag{3.4}$$

see also Fig. 3.1. Following Refs. [68, 69], I set $r_0 = 0.7$, $R = 0.3$, and $K = 40$ such that the maximal bond length is the unit length. All *non-bonded* monomers interact via the 12-6 Lennard-Jones (LJ) potential

$$V_{\mathrm{LJ}}(r) = 4\epsilon \left[(\sigma/r)^{12} - (\sigma/r)^6\right], \tag{3.5}$$

with $\epsilon = 1$ and $\sigma = r_0/2^{1/6}$. The latter is chosen such that the potential minimum coincides with the minimum of the FENE potential. In addition, the Lennard-Jones potential is cutoff above a cutoff-radius r_c which allows to employ a domain decomposition. In this case, the periodic box is divided in equally sized domains of minimal linear length r_c. This allows to reduce the computation of (partially negligible) contributions to the (relevant) monomers in the same or neighboring domains only, of course taking periodic boundary condition into account. In order to be consistent with the aforementioned literature, the Lennard-Jones potential is cutoff at $r_c = 2.5\sigma$ (compare also Sec. 2.1.2) such that

$$V^*_{\mathrm{LJ}}(r) = \begin{cases} V_{\mathrm{LJ}}(r) - V_{\mathrm{LJ}}(r_c) & r < r_c \\ 0 & \text{else} \end{cases}. \tag{3.6}$$

This self interaction leads to a collapse transition of a single polymer. These polymer models are also referred to as Θ-polymers. Adding polymers to the system, the mutual attraction eventually leads to aggregation.

Most results are obtained again using parallel multicanonical simulations. The usual praxis includes parallel-tempering pre-runs only for the smallest system sizes. This allows to extrapolate the lower and upper energy-ranges for the multicanonical simulation of larger system sizes. In general, I perform between 1.28 and 2.56 million measurements with $1 - 100$ sweeps in between. The sweeps are of the order of $N \times M$ updates, drawn randomly from a set of update moves including single-monomer displacement, bond-rotation, polymer translation, and the more sophisticated double-bridging move. The latter move proved to be important especially in the low-temperature aggregated phase, where the flexible polymers form a spherical, entangled object. Making use of the generalized ensemble approach, I employed adapted variable update ranges with bias correction [70] to optimize the acceptance rates in every part of the energy landscape. For details to the implementation of updates and a test of proper sets of update moves see Sec. 5.7.

In order to validate the implemented model and multicanonical method, I consider an additional, independent implementation of the model. There, I use straight-forward, non-optimized functions for the energy calculation, consider only single monomer displacements and employ

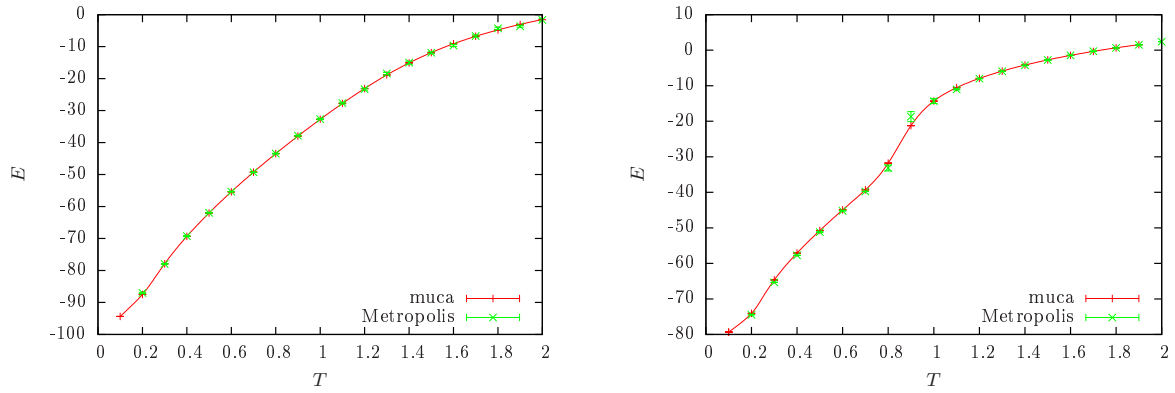

Figure 3.4.: Qualitative comparison of the canonical energy from parallel multicanonical simulations and independent Metropolis simulations. (left) Single flexible polymer with $N = 30$ and (right) aggregation of a 2×13 flexible polymer system.

the standard Metropolis algorithm. The simulation started at high temperature and after a thermalization period, 20 000 measurements with 52 000 sweeps were performed for each temperature. Figure 3.4 shows both approaches for a single polymer of length $N = 30$ (left) undergoing a collapse transition and for 2 polymers of length $N = 13$ (right) with the corresponding aggregation transition. The errors for the multicanonical data are obtained by jackknife error analysis and Metropolis error bars are the statistical errors including the integrated autocorrelation time, for details see e.g. Refs [6, 71]. Both examples show that the Metropolis averages scatter around the estimates from the multicanonical data, consistent within the error bars. In the vicinity of the transition the integrated autocorrelation time increases, in case of the first-order like aggregation transition this effect is very prominent. The corresponding fewer independent contributions to the Metropolis estimates explain the larger deviations near $T = 0.9$. In addition, the employed domain decomposition was tested on exemplary cases of $M \times 20$ polymers in a 60^3 box with periodic boundary conditions. Therefore, I use Metropolis simulations at $T = 1$ with 1 million sweeps in order to directly compare the time of energy calculation. Starting at $M \approx 4$ polymers, the domain decomposition showed a speedup while the time series remained identical.

3.2. Effect of inter- and intra-polymer interaction

In order to discuss the interplay of inter- and intra-polymer interaction, I will focus on the example of two lattice polymers of length $N = 9$. As shown above, this case is still treatable by exact enumerations, here on a cubic lattice of size 35^3 yielding a monomer density $\rho_\mathrm{m} \approx 4 \times 10^{-4}$. The Hamiltonian in Eq. (3.3) may be rewritten in terms of the coupling ratio $\epsilon' = \epsilon_\mathrm{inter}/\epsilon_\mathrm{intra}$ between inter- and intra-polymer interaction:

$$\mathcal{H} = -\epsilon_\mathrm{intra} \left(n_\mathrm{intra} + \epsilon' n_\mathrm{inter} \right) . \tag{3.7}$$

Setting $\epsilon_\mathrm{intra} = 1$ in practice coincides to a rescaled temperature $T' = T/\epsilon_\mathrm{intra}$ and allows to only adjust a single parameter ϵ'. Moreover, I perform exact enumerations measuring both the number of intra- and inter-polymer contacts, see Sec. 5.2. This yields a two-dimensional

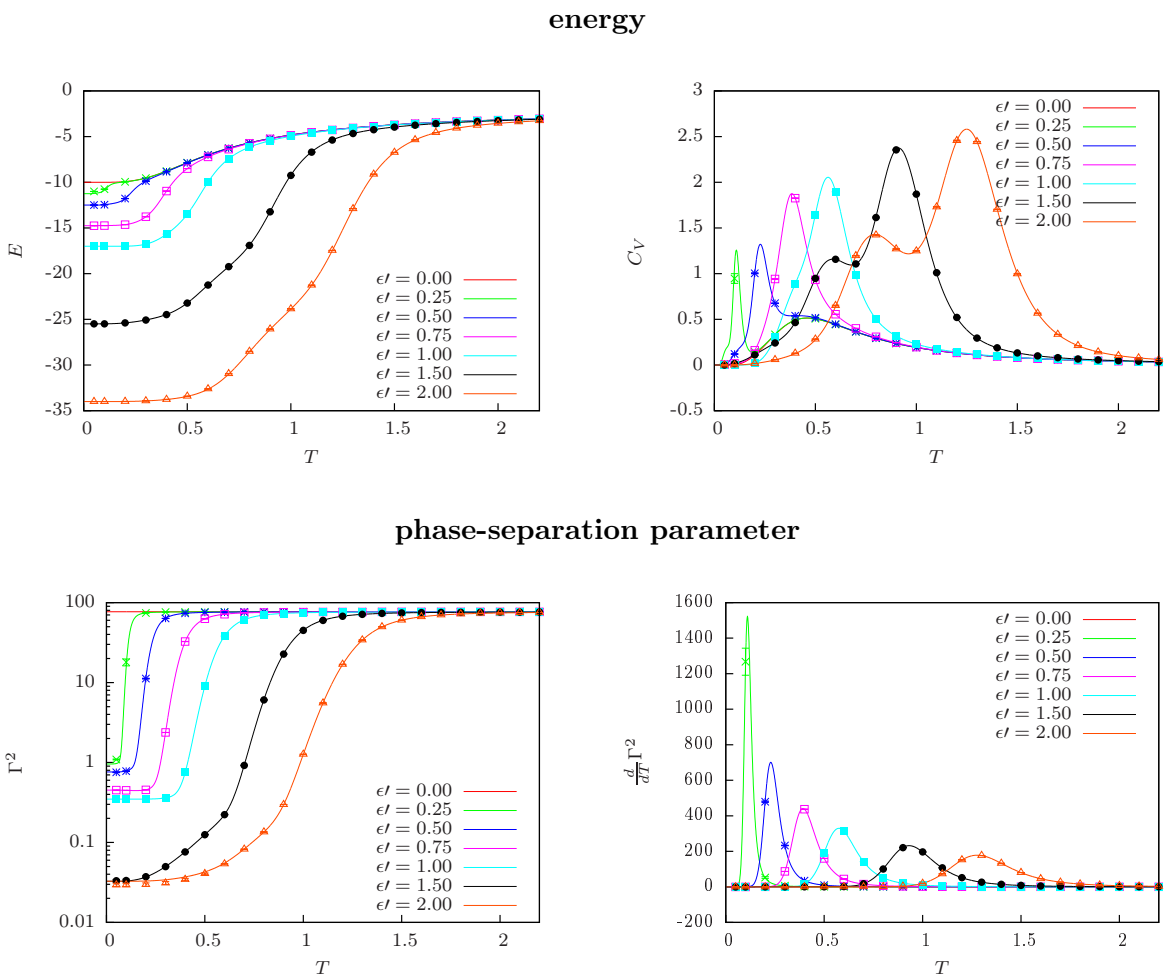

Figure 3.5.: Canonical expectation values (solid lines) of the energy and phase separation parameter (left) and their thermal derivatives (right) from exact enumerations of a 2×9 polymer system for selected coupling ratios ϵ'. Data symbols with error bars in the corresponding color are obtained from parallel multicanonical simulations.

density of states and allows to reweight to canonical expectation values for any ϵ'. This includes the case $\epsilon' = 1$ commonly considered in this thesis. In addition, multicanonical simulations of these reference systems allow to compare the application of non-integer coupling ratios in multicanonical simulations and yield exemplary conformations in the occurring structural phases. The canonical expectation values in the following are always combined with estimates from multicanonical simulations.

Figure 3.5 (top) shows the canonical expectation value of the system energy and its derivative the specific heat for two polymers of length $N = 9$ with selected coupling ratios ϵ'. The case $\epsilon' = 0$ is comparable to the energy curve of a single polymer, showing only a collapse signature in the specific heat. In this case, the polymers do not interact energetically and behave as isolated polymers, if the system is sufficiently dilute. Increasing the inter-polymer interaction strength slightly leads to an energy decrease at small temperatures. The large peak in the specific heat

shows that the associated aggregation transition temperature is below the collapse transition temperature (broad peak). For comparable interaction strengths ($\epsilon' \approx 1$) the collapse and aggregation transition are almost indistinguishable. However, for inter-polymer interactions sufficiently larger than intra-polymer interactions ($\epsilon' > 1$) it is clearly visible in both energy and specific heat that there are again two separate transitions occurring. By the height of the peaks it may be argued that for $\epsilon' > 1$ at large temperatures the aggregation occurs, while the peak at lower temperatures needs further insight.

The phase separation parameter Γ^2 was introduced to describe the aggregation transition and is shown in Fig. 3.5 (bottom) together with its thermal derivative. For $\epsilon' = 0$, the phase separation parameter remains large, confirming that no aggregation occurs. With increasing ϵ' a clear aggregation transition is observable: for small temperatures, the polymers are close to each other resulting in a small Γ^2; and for large temperatures, independent separated polymers result in a large Γ^2. The transition is accompanied by a peak in the thermal derivative of the phase separation parameter. Interesting to notice is that Γ^2 is smaller for larger ϵ'. This indicates that the resulting structures for large ϵ' are more compact in the sense that the centers of mass are closer together. Similar observations have been made for stiff polymer bundles before [65].

In order to describe the involved collapse transition in more detail, the number of intra-polymer contacts $n_{\text{intra}} = n_{\text{i}}$ is a good observable. The expectation value of the number of intra-polymer contacts is shown with its thermal derivative in Fig. 3.6 (top). The case of isolated polymers, $\epsilon' = 0$, is again a good reference for the single-polymer collapse. With decreasing temperature, the number of intra-polymer contacts increases. This is a continuous transition and the corresponding transition temperature may be identified by the negative peak in the thermal derivative of the number of contacts. With increasing ϵ', the number of intra-polymer contacts in the low-temperature state decrease. For small ϵ' this again signals an aggregation transition (positive peak in the derivative) at a lower temperature than the collapse transition. With increasing ϵ' the negative collapse peak in the derivative becomes weaker up to a point, where a peak is technically there, but meaningless with respect to structural transitions. This is due to the continuous nature of the single-polymer collapse transition. As long as the polymers are isolated in the high-temperature regime, the number of intra-polymer contacts will gradually increase with decreasing temperature. For sufficiently large $\epsilon' \gtrsim 1$ the aggregation occurs at temperatures at or above the single-polymer collapse temperature. It seems that the aggregation dominates over the (initiated) collapse, decreasing the intra-polymer contacts again and what remains is a small negative peak in the derivative without significance. For $\epsilon' \approx 1$, the initial decrease in n_{intra} is followed by an increase when lowering the temperature. This changes for larger ϵ', where the initial decrease is followed by a second, stronger decrease down to 0 contacts at low temperature, seen by two positive peaks in the derivative.

The expectation value for the number of inter-polymer contacts $n_{\text{inter}} = n_{\text{o}}$ is shown in Fig. 3.6 (bottom) together with its thermal derivative. As for the phase separation parameter, the case $\epsilon' = 0$ shows constant behavior. With increasing ϵ', polymer contacts in the low-temperature conformations increase, consistent with the decrease in the number of intra-polymer contacts. For large temperatures, the separated polymers are practically not in contact, i.e., $n_{\text{inter}} = 0$. At the aggregation transition, n_{inter} increases, also shown by a large negative peak in the derivative. For $\epsilon' \approx 1$, n_{inter} assumes a plateau, or even slightly decreases at lower temperatures. For large ϵ', there is a second peak visible, showing that the loss of intra-polymer contacts was in favor of an additional gain of inter-polymer contacts.

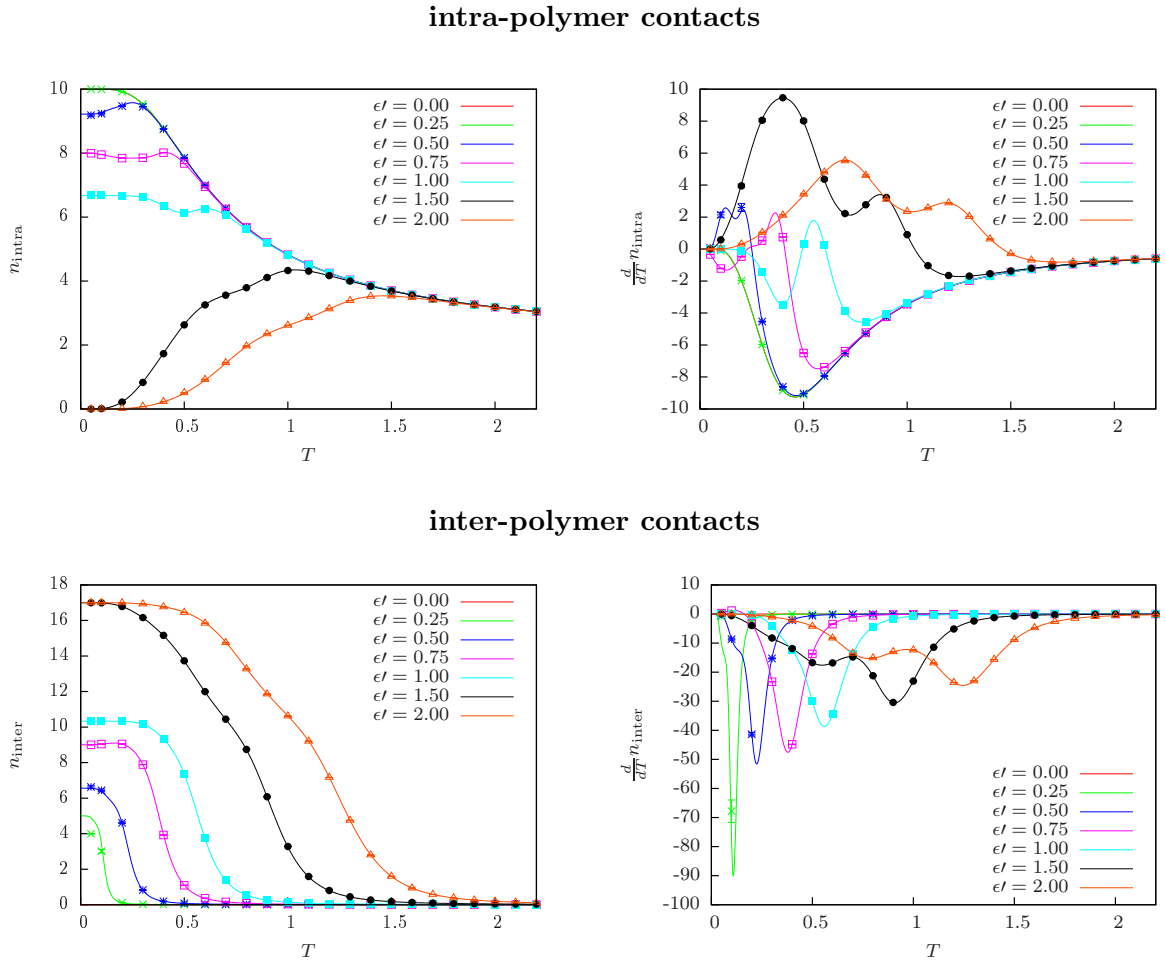

Figure 3.6.: Same as Fig. 3.5 for the intra- and inter-polymer contacts (left) and their thermal derivatives (right).

A complete overview for the finite 2×9 system is provided in Fig. 3.7. The structural phase diagram is obtained from a surface plot of the specific heat and the transition points obtained from local positive maxima and negative minima, as discussed above. This includes the artificial "transition" signatures mentioned before. Therefore, a sketch of the structural phase diagram is given in Fig. 3.7 (right). Combined with the previous discussion, four different regions may be identified already for this small system. For large temperatures, the two polymers are separated and each polymer is in the extended state (SE). Reducing the temperature for small ϵ', the individual polymers undergo a collapse transition and thus become separated globules (SG). The aggregation transition spans the diagram as a diagonal, below which both polymers are close to each other. Directly below the aggregation transition, the inter- and intra-polymer interactions are competing and one finds the two polymers attached to each other, with a remaining self-interaction (A). However, with increasing ϵ', the inter-polymer interactions dominate at low temperatures and an additional structure occurs, where inter-polymer contacts dominate in an "entangled" conformation (A'). Exemplary conformations from multicanonical simulations are provided below the structural phase diagrams. The dif-

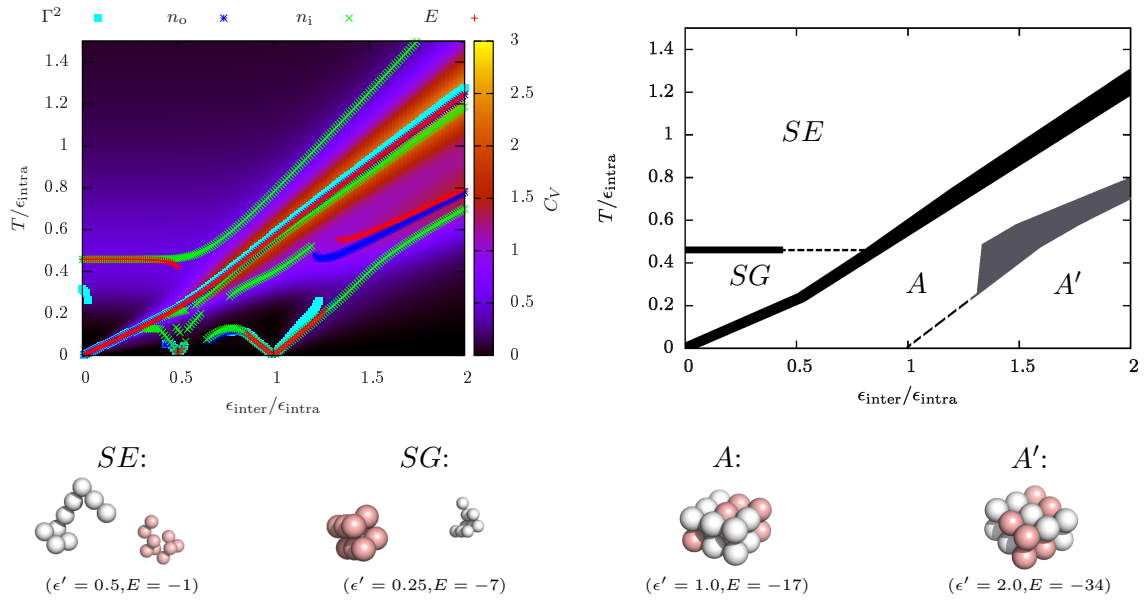

Figure 3.7.: Structural phase diagram of two flexible lattice polymers ($N = 9$,$L = 35$) with varying coupling ratio ϵ' from exact enumerations. (left) The surface plot shows the specific heat and the symbols mark local positive maxima and negative minima in the thermal derivatives of the corresponding observables. (right) A simplified sketch with the identifiable structural phases. S and A refer to separated and aggregated conformations; E and G stand for extended and globular single-polymers; and A' conformations are entangled with maximizing inter-polymer contacts. Examples from multicanonical simulations are provided below the figures.

ference between conformations from A and A' are subtle, but present in terms of self-contacts. Conformations in A may be considered as globules that stick together. In some cases, this reminds of jigsaw pieces, where the individual polymers rearranged in such a way to fit to each other. In contrary, conformations in A' have a decreasing number of self-interactions, trying to maximize inter-polymer contacts. This leads to entangled conformations, where the polymers mutually "wrap" around each other. On close sight, this may be identified in the presented examples.

An interesting observation is the similarity to homopolymer adsorption on flat, flexible and patterned surfaces showing very similar phase diagrams [54, 72–75]. This is not unexpected but an artifact of considering only two polymers. The equivalence between bundling of two semiflexible polymers and adsorption of a semiflexible polymer to a surface has been exploited before, e.g., in Ref. [76]. In this case, the transition $A \to A'$ may be identified with the transition into single layers on a flexible surface [73] or the recognition transition onto a patterned surface [75]. The latter case could serve as an intuitive picture, if both polymers are considered to "recognize" the patterned one-dimensional surface of the respective partner, which is energetically favored.

The presented results show a qualitative picture of the effect of inter- and intra-polymer coupling on the structures of homopolymer aggregates using a finite lattice polymer model

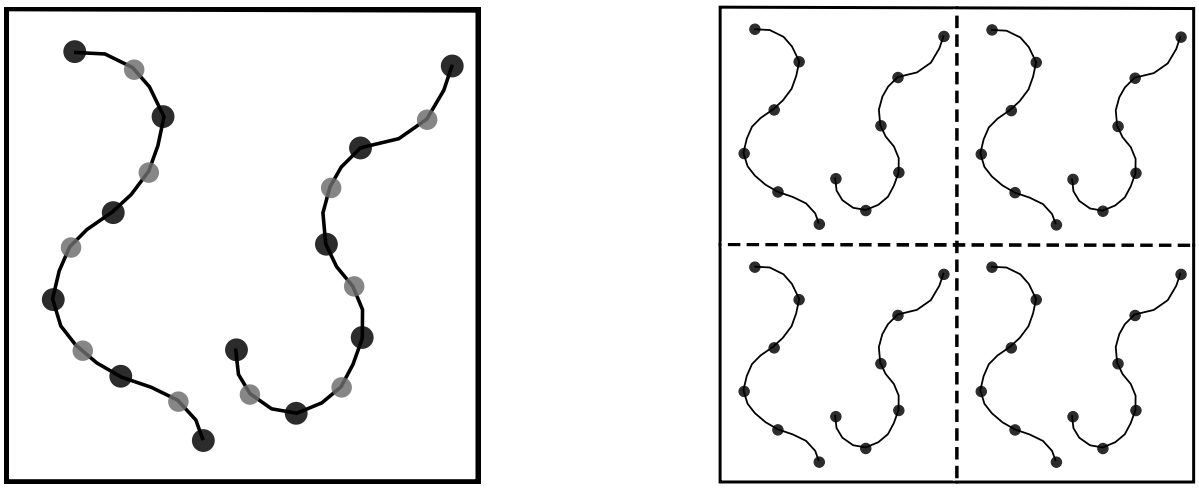

Figure 3.8.: Sketch of different approaches to the thermodynamic limit for polymer aggregation: fixing the polymer number and increasing polymer length in terms of monomers (left) or fixing the polymer length and increasing the polymer number (right).

which, however, may be treated with exact enumerations. This serves as a starting point for a more detailed investigation of occurring structures both in the limit of two longer chains ($M = 2$, $N \to \infty$) and many short chains ($M \to \infty$, N fixed) for selected coupling ratios ϵ'. This is currently investigated by a Master student in our group, Benjamin Schott [77]. The example $\epsilon' = 1$ is discussed for both limits in Sec. 3.3.2 and Sec. 3.4 respectively.

3.3. Finite-size scaling limits in polymer aggregation

For polymer solutions, the thermodynamic limit may be reached in two ways: either increasing the number of polymers or increasing the length of a single polymer (see e.g. Ref. [63]). The usual description of universal properties for generic polymer models is in the limit of infinitely long (isolated) polymer chains $N \to \infty$. It is this limit, in which for example discrete and continuous self-avoiding walks show the same predicted scaling behavior. Here, a mapping to ordinary finite-size scaling of spin systems is possible, considering the $O(n)$ model with $n \to 0$ [3]. For a single (isolated) polymer, translational entropy may be neglected, if a homogeneous solution is assumed. Thus, single polymers may be considered in open systems or periodic boundaries with $L \gg Nr_0$ and the monomer density does not play a role. For the current bead-spring model with self-interaction (Θ-polymer), the scaling of the resulting collapse transition is revisited in Sec. 3.3.1.

Adding already a second Θ-polymer leads to additional length-scales which introduce difficulties in the definition of a suitable finite-size scaling limit. Now density becomes important because separate polymers need to find each other. A detailed discussion of the effect of density is given in Sec. 3.5. Sticking to the limit $N \to \infty$, a fixed monomer density seems to be an unsuitable choice for a proper finite-size scaling of polymer aggregation. This is because in this limit the polymers grow linearly, while the linear system size L only grows with $N^{1/3}$. This eventually leads to the point $N \gg L$, where the polymer interacts with itself across periodic boundaries or is influenced by steric boundaries. Another possibility is to define a fixed "renormalized" polymer density $\rho = M/L'^3$, where $L' = L/N$ sets the length scale. In

this limit, the polymers may be considered to become more resolved going to the continuum limit. An example is shown as a sketch in Fig. 3.8 (left) where L' is fixed and the polymer length is increased by adding additional beads. However, considering Θ-polymers, the interplay with short-range Lennard-Jones interaction becomes non-trivial. This is because the polymer-polymer interaction length scale remains on the order of the intra-polymer interaction length set by the bond length. Compared to the system size or polymer length, the polymer-polymer interaction length scale thus decreases. In the limit of infinitely long (or infinitely resolved) chains, the inter-polymer interaction range would approach zero. This is exemplary discussed in Sec. 3.3.2.

Another possibility to reach the thermodynamic limit is to consider the limit of infinitely many polymers $M \to \infty$ of fixed length N. Increasing the number of polymers linear with the volume fixes the polymer density $\rho = M/V$. See for an example Fig. 3.8 (right), where for each volume element the same number of polymers are added. This is a proper finite-size scaling limit, where the length scale of polymer-polymer interaction compared to the single-polymer length remains the same for all system sizes. Also, at least in the homogeneous phase, the volume fraction is locally conserved with increasing system size. In this limit, the length of the polymer N may be considered as a chemical property that is fixed in the finite-size scaling. This is content of Sec. 3.4 and is analogue to the definition of finite-size scaling for particle gas, considering polymers as "extended" particles.

3.3.1. The collapse transition revisited

In the past 10 years, there has been an interesting discussion in the literature about the scaling of the collapse transition temperature of a single flexible homopolymer. Among others, this problem has been discussed using lattice models [63, 66] as well as continuous bead-spring models [78–80]. One usually expects a scaling motivated from polymer solutions in the Flory-Huggins mean-field theory [2]. This leads to the fit ansatz

$$T_\Theta(N) - T_\Theta = -\frac{a_1}{\sqrt{N}} + \frac{a_2}{N}, \tag{3.8}$$

where $T_\Theta = \lim_{N\to\infty} T_\Theta(N)$. This has shown to correctly describe the finite-size corrections to the collapse transition temperature for interacting self-avoiding walks [66], the bond-fluctuation model [63] and continuous bead-spring polymers [80]. However, in Ref. [78, 79] the authors claim to observe mean-field predicted logarithmic corrections leading to

$$T_\Theta(N) - T_\Theta = \frac{A}{\sqrt{N}(\ln N)^{7/11}}. \tag{3.9}$$

The aforementioned contributions [63, 66, 80] have shown that this is rather unlikely. One interesting observation is that the numerical verification of Eq. (3.9) relied on numerical derivatives of sampled densities of states including smoothening. This indicates that here some caution is necessary.

I will focus on the first ansatz Eq. (3.8) in order to characterize and compare two slightly different models of a bead-spring polymer: a bead-spring polymer with non-bonded Lennard-Jones (LJ) interaction (described in Sec. 3.1.2) and a bead-spring polymer with additional bonded Lennard-Jones interaction, e.g., applied in Ref. [69, 78, 79] with the same parameterization. Attempted fits to the second ansatz did not show satisfying results. For large tem-

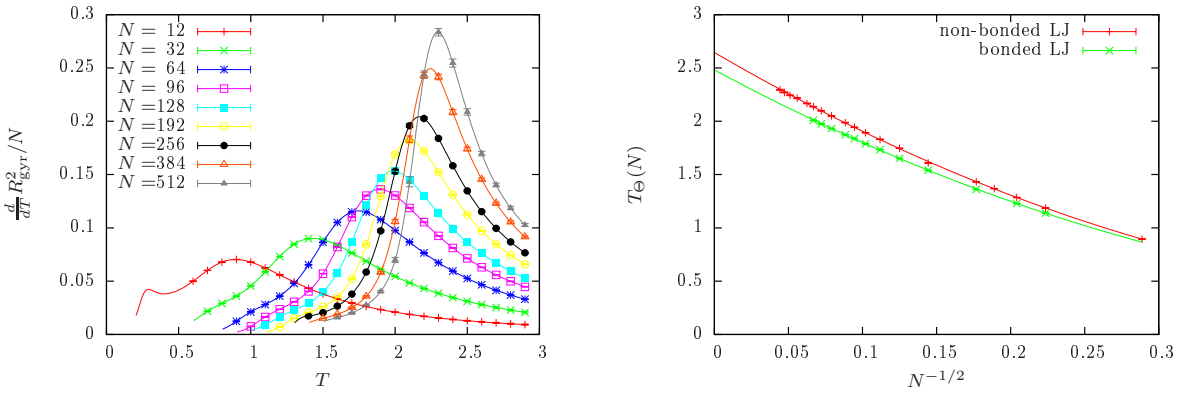

Figure 3.9.: (left) The canonical temperature derivative of the squared radius of gyration for the bead-spring polymer with non-bonded LJ-interaction. (right) Finite-size scaling of the collapse transition temperature. The data is nicely described by Eq. (3.8) (solid lines).

peratures, the polymer is entropy dominated and the radius of gyration scales as self-avoiding walk with increasing chain length. However, with decreasing temperature the short-range attraction becomes more relevant and the polymer undergoes a continuous collapse transition, forming a more compact globule. This is reflected in a decrease of the radius of gyration and the maximal slope may be associated to the collapse transition temperature. A similar definition via the change in energy or the maximum/shoulder of the specific heat is possible and will yield a slightly different finite-size temperature. However, in the thermodynamic limit all proper definitions should yield the same estimate. Here, $T_\Theta(N)$ is obtained from the peak of the squared radius of gyration R^2_{gyr}. The canonical curves for the non-bonded LJ case is shown in Fig. 3.9 (left). For larger system sizes, the collapse transition shifts to larger temperatures and the peak height increases.

The shift in the collapse transition temperature for finite systems is shown in Fig. 3.9 (right) for the two considered models. The errors are obtained by jackknife error analysis and the data points are plotted against the expected $N^{-1/2}$ scaling behavior. In both cases, the additional higher-order corrections are clearly visible by a non-linear behavior in the plot variables. However, the fit of Eq. (3.8) describes the data points well. For details of the fits and a comparison to a differently parameterized bead-spring polymer [80] see Table 3.2. The differences for both thermodynamic limit T_Θ and finite-size corrections are due to (slightly) different energy (and corresponding temperature) scales from the microscopic details of the models. Consider, for example, the non-bonded LJ interaction and the additional bonded LJ interaction. In the

Table 3.2.: Fit results to the scaling Ansatz [Eq. (3.8)] of the single-polymer collapse transition temperature for three bead-spring models.

Model	T_Θ	a_1	a_2	χ^2
non-bonded LJ	2.645(4)	8.10(6)	7.1(2)	≈ 1.98
bonded LJ	2.481(8)	7.5(1)	6.4(4)	≈ 1.22
Ref. [80]	3.176(4)	11.56	12.77	

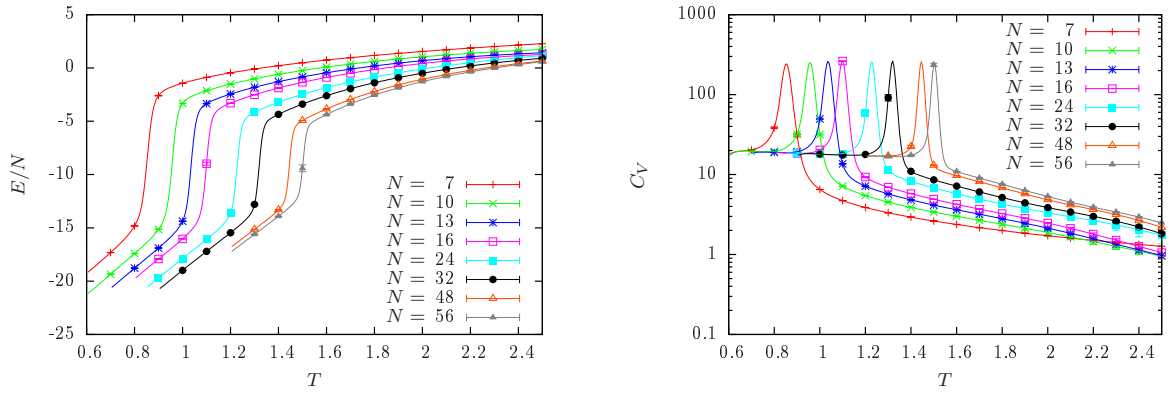

Figure 3.10.: Exemplary canonical estimates of the energy (left) and the specific heat (right) for a fixed number of polymers $M = 8$ with variable length N at fixed density $\rho' = M/L'^3 = 0.1$. The ratio $L' = L/N$ is kept fixed.

latter case, bonded monomers yield an energy contribution even at large temperatures because the FENE potential keeps bonded monomers at close distance. This is, however, no constant energy term and results in a lower transition temperature. The model in Ref. [80] also includes bonded LJ interaction but uses different parameters which explains a different T_Θ. However, the important result is that all models follow the same finite-size behavior and are thus expected to describe the same generic properties. For the influence of interaction range on the interplay of collapse and freezing transition see also Ref. [81], where the present model is clearly in the second-order like regime.

3.3.2. Aggregation of flexible polymers in the limit $N \to \infty$

As mentioned in the beginning of this section, one possible approach to the thermodynamic limit could be a fixed polymer number M with increasing polymer length N (in number of monomers). Then, it seems unavoidable to couple the polymer length to the linear system size, thus considering the limit of a continuous polymer string. This limit is known from a worm-like chain perspective on single semiflexible polymers or polymer bundles with cross-linkers. The considered bead-spring model, with equal intra- and inter-polymer interaction ($\epsilon' = 1$ in Sec. 3.2), seems to be inadequate for this limit. This is because the interaction length scale of both intra- and inter-polymer Lennard-Jones interaction is coupled to the bond length. In the limit of infinitely resolved single polymers, the bond length tends to zero. For the intra-polymer interaction, and the accompanying collapse transition in Sec. 3.3.1, this is effectively renormalized by the diverging number of monomers. Considering, however, a fixed $L' = L/N$ for several polymers in the limit $N \to \infty$ means that the polymers become infinitely thin and the inter-polymer interaction length vanishes compared to the system size. As a consequence, the polymers do not "see" each other very well compared to the huge entropy of the solution but once they find each other the energy gain is enormous.

Exemplary canonical estimates for fixed $M = 8$ and increasing N are shown in Fig. 3.10: the average energy (left) and the specific heat (right), normalized to the number of monomers, i.e., $C_V = k_B\beta^2 \left(\langle E^2\rangle - \langle E\rangle^2\right)/N$. The transition temperature is obtained from the largest peak in the specific heat. At the aggregation transition, the energy shows a sharp jump for

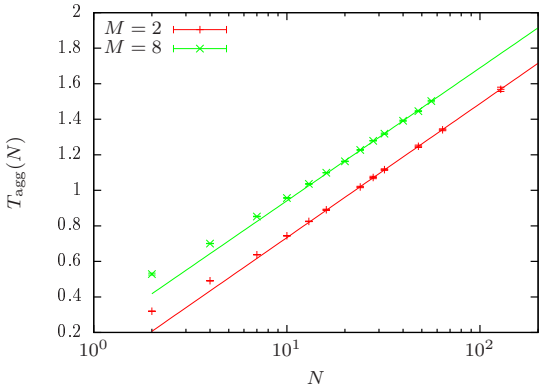

Figure 3.11.: Scaling of the aggregation transition temperature for fixed M and increasing N. The rescaled density is $M/L'^3 = 0.1$, such that $L' = L/N$ is fixed for each curve.

all system sizes between a high-energy regime at high temperatures and low-energy regime at low temperatures. Below the aggregation temperature, the specific heat curves coincide for all polymer sizes while qualitative deviations at high temperatures are noticeable. Interesting to notice is that in this limit the normalized energy difference at the transition seems to decrease and the peak in the specific heat does not increase with the linear size of the polymers. The first observation may be understood considering that the polymers start to collapse around the same temperature range, compare also Sec. 3.3.1. Therefore, the increase in N also increases the possible number of intra-polymer contacts, which already reduces the average energy above the aggregation transition in the fragmented regime.

In this limit, the canonical observables do not show the expected first-order like characteristics of a decreasing width and increasing height of the specific heat peak, or an increasing sharpness of the energy jump. What can be said is that the aggregation transition temperature increases with polymer length. This suggests that the energetic gain from the increase in possible inter-polymer contacts dominates over the increase in entropy due to the rescaled interaction length scale. A heuristic functional dependence on the polymer length $T = a + b \ln N$ is shown in Fig. 3.11. For both exemplary cases $M = 2$ and $M = 8$, corresponding fits to the data with $M \gtrsim 20$ yield slopes $b \approx 1/3$ with decent quality of fits (reduced χ^2), namely $\chi^2 \approx 0.1$ and $\chi^2 \approx 2.7$, respectively. Alternatively, also power-law fits yield adequate results. This suggests that for a fixed number of coarse-grained Θ-polymers, the limit $N \to \infty$ does not describe a proper phase transition in the usual sense. On the contrary, in the limit of infinitely long polymers, the aggregation temperature diverges. For the considered model, it may be thus argued that polymers need to have a finite length as well as finite interaction strength and length scale. Noteworthy, this is consistent with reality where polymers and proteins are always finite. The two considered cases are, however, special in the sense that M is very small.

Figure 3.12.: Snapshots of a separated (left) and aggregated (right) conformation for $M = 48$ bead-spring polymers of length $N = 13$.

3.4. Finite-size effects and scaling regimes in the limit $M \to \infty$

As mentioned before, equilibrium aggregation describes the transition between a separated phase of individual polymers and a homogeneous aggregated phase in which a fraction of polymers are condensed in a single macroscopic object. Figure 3.12 shows two corresponding snapshots from a multicanonical simulation with $M = 48$ bead-spring polymers of length $N = 13$. The equilibrium description of this problem results in homogeneous aggregation, where in non-equilibrium considerations also heterogeneous aggregation (several aggregates) occurs due to increasing relaxation times. The common notion of phases in the thermodynamic limit is recovered when considering the limit of increasing polymer number $M \to \infty$ at fixed polymer length N and polymer density $\rho = M/V$. Then, the polymer length may be considered as a system property. In fact, for a fixed N one may equivalently consider the monomer density $\rho_\mathrm{m} = \rho N$. This limit is quite similar to the description of particle gas condensation at fixed density, where the polymers may be considered to be extended "particle-like" objects. Consequently, one would expect the same qualitative finite-size scaling behavior as in Sec. 2.3: above the aggregation transition the polymers are independent of each other, individually exploring the available space, and below the aggregation transition there exists a mixed phase of the single macroscopic aggregate in equilibrium with a surrounding polymer solution. I will present results for $N = 13$ lattice and bead-spring polymers at fixed monomer density $\rho_\mathrm{m} = 0.01$ with a focus on the lattice polymers due to numeric reasons. In case of lattice polymers, the density is adjusted approximately. For a description of the models see Sec. 3.1. All main results are obtained with parallel multicanonical simulations using up to 256 cores (see Sec. 5.4), recording in total 1.28-2.56 million measurements. Relevant data presented in the plots are listed in tables in the appendix. The update moves for the lattice case include pivot, corner flip, slithering snake, translation, and double bridge moves. In the case of bead-spring polymers, I considered single-monomer displacement, bond rotation, translation, and double bridge moves. For details see Sec. 5.7.

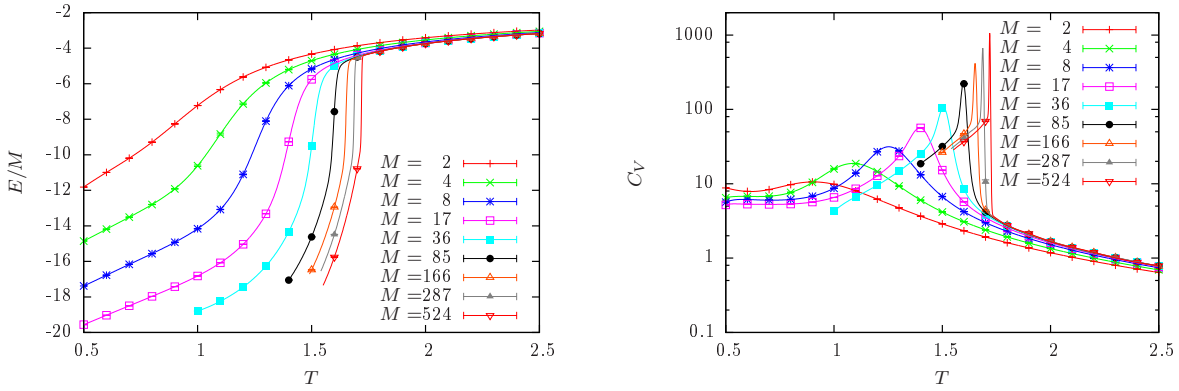

Figure 3.13.: Canonical estimates of the average energy (left) and the specific heat (right) per polymer for a lattice polymer system with $N = 13$ and $\rho_{\mathrm{m}} = 0.01$.

The aim of this section is to compare flexible homopolymer aggregation to particle gas condensation using finite-size scaling, however mostly in an appearing intermediate scaling regime. I will consider observables analogue to the particle condensation discussion in Chap. 2, namely the aggregation temperature, the rounding of the transition and the cluster size at aggregation. Additionally, I will discuss the finite-size scaling in the microcanonical ensemble.

The general concept of finite-size scaling is to use small system sizes in order to make predictions about the thermodynamic limit, or systems of infinite size, and to describe how certain physical properties change under variation of system size. For second-order phase transitions, this allows to study critical phenomena including critical exponents and universality, see also the discussion of the collapse transition in Sec. 3.3.1. The aggregation transition is usually identified as a first-order like transition, or pseudo phase transition, since the usual system sizes and constituents are finite by nature. Considering the formal similarity to particle condensation and the competition between entropy maximization in the separated phase and energy minimization in the aggregate, it seems justified to reapply the notion of a first-order phase transition in the limit of infinite polymer number. In order to get an impression, Fig. 3.13 shows the average polymer energy and specific heat, normalized to the number of polymers, i.e., $C_V = k_B\beta^2 \left(\langle E^2\rangle - \langle E\rangle^2\right)/M$, for lattice polymers. Compared to the case $N \to \infty$ at fixed $M \neq 1$ in Sec. 3.3.2, here the specific heat peak grows in height and shrinks in width. This already gives a hint that the limit $M \to \infty$ allows for a proper definition of a thermodynamic limit for dilute polymer aggregation, analogue to particle condensation.

Before comparing the scaling of relevant observables, I briefly discuss the finite-size scaling expectations. Building on Sec. 2.3, and considering the similarity of dilute particle condensation to dilute Θ-polymer aggregation, one may expect that the aggregation transition separates a homogenous and inhomogeneous (or mixed) phase. Considering that the mixed phase is supposed to thus consist of a single macroscopic aggregate with surrounding "gaseous"/solute polymers allows to identify the radius R of the aggregate as the *relevant linear length scale* [42, 45, 46]. Then the surface $\left(R^{(d-1)/d}\right)$ and volume $\left(R^d\right)$ contributions to finite-size corrections compete, which would give rise to a finite-size scaling correction of the order R^{-1}.

For flexible polymers it may be assumed that the aggregate shape is spherical, just like a droplet of monomers. Thus, from the experience with particle condensation in Sec. 2.3, one may expect two regimes: a large-system scaling behavior, where $R \sim M^{1/4}$, and an intermedi-

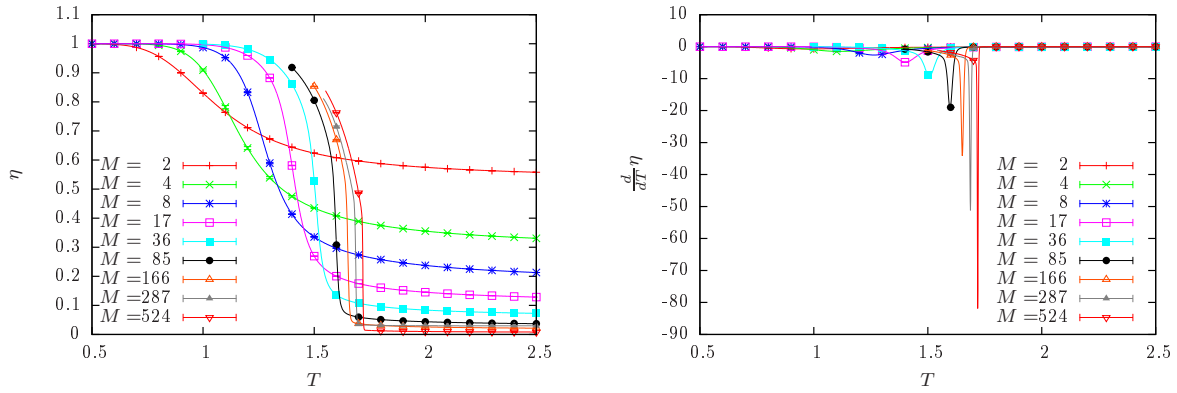

Figure 3.14.: Canonical estimates of the fraction η of monomers in the largest cluster (left) and its thermal derivative (right) for a lattice polymer system with $N = 13$ and $\rho_{\mathrm{m}} = 0.01$.

ate effective regime, where $R \sim M^{1/3}$. The large-system condensation-like expectation is based on the observation that in the gaseous (separated) phase the individual polymers maximize entropy by homogeneously distributing in space, while in the aggregated phase the flexible polymers minimize energy forming a single aggregate in equilibrium with unattached polymers. Following the arguments from condensation, comparing free-energy contributions from fluctuations and surface, leads to a macroscopic aggregate including only a fraction $\eta \sim M^{1/4}$ of polymers. With increasing system size, the aggregate would grow slower than the total number of polymers, reflected in a linear extension $R \sim M^{1/4}$. In fact, the aggregate of several polymers is similar to the collapsed state of a single, long polymer as well as the condensate of non-bonded monomers (gas) [69]. A non-trivial difference is, however, the internal energy and entropy of the individual polymers in the separated phase.

In addition, an effective intermediate scaling regime may occur similar to particle condensation. It may be argued that for a *mesoscopic*, or intermediate, number of polymers a large fraction of polymers will contribute to the aggregate. This is a reasonable assumption as can be seen in Fig. 3.14: the fraction η of lattice polymers ($N = 13$) in the largest cluster shows a jump from close to zero at high temperatures to a non-zero fraction at the transition temperature. Lowering the temperature further, the fraction of polymers in the cluster/aggregate increases. For most system sizes, $\eta \gtrsim 0.5$ in the aggregate below the transition temperature. This would justify an extended intermediate regime. Assuming a spherically shaped aggregate, the relevant length scale is the corresponding radius. For a large fraction of uniform polymers (fixed N) in the aggregate this would lead to $R \sim M^{1/3}$. In order to investigate the occurring scaling regimes, I start by recapturing the expected scaling form of each observable generalized from the arguments and results of particle condensation (see Chap. 2).

3.4.1. Transition temperature and rounding

For the aggregation temperature one expects a finite-size shift of the order of the inverse relevant linear system size. This means

$$T_{\mathrm{agg}} = T_0 + aM^{-p}, \tag{3.10}$$

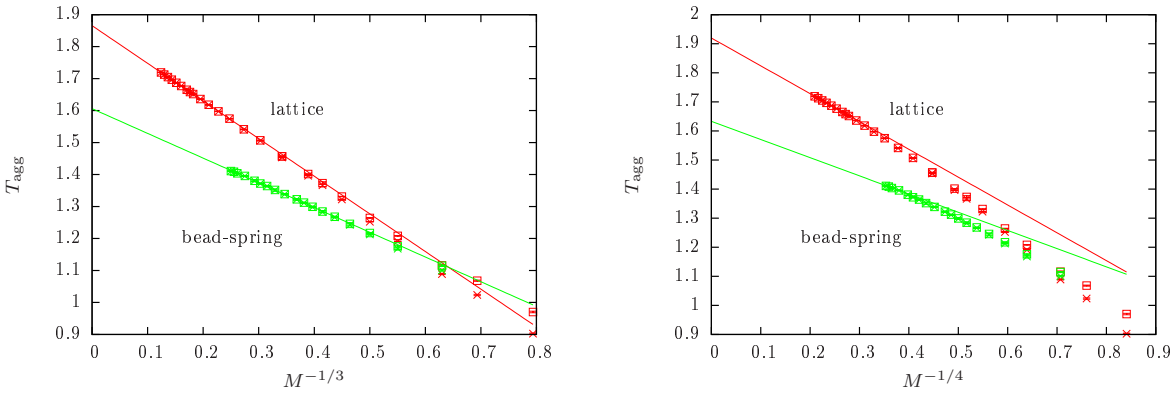

Figure 3.15.: Finite-size scaling of the aggregation temperature for $N = 13$ flexible lattice (red) and bead-spring (green) polymers with $\rho_m = 0.01$. The x-axis is scaled in order to show the expected scaling of the intermediate (left) and large-system (right) regime. The transition temperatures are obtained from the peaks of the specific heat (x) and thermal derivative of η (box). The solid lines represent the best fits according to the expected scaling.

where $p = \{\frac{1}{3}, \frac{1}{4}\}$ is the expected exponent for the effective intermediate or large-system regime. Figure 3.15 shows the aggregation temperature plotted against $M^{-1/3}$ (left) comparing to the predicted intermediate behavior and against $M^{-1/4}$ (right) comparing to the expected large-system behavior. Both plots show results for lattice (red) and bead-spring (green) polymers. Indeed, the numerical data show that the change in aggregation temperature decreases with increasing system size, justifying the ansatz in Eq. 3.10. It can be seen that for a large range of M values, the $M^{-1/3}$ rescaling shows a linear behavior. This indicates an effective intermediate scaling regime. However, for larger system sizes also the $M^{-1/4}$ rescaling indicates a linear behavior. In both cases, reasonable fits with goodness of fit parameter $Q \gtrsim 0.5$ are achievable, shown by the solid lines in the figure and listed in Table 3.3. A possible implementation of the goodness of fit parameter Q for gnuplot is given in Ref. [82]. The achievable data for the lattice gas exceeds the bead-spring case about an order of magnitude. However, it may be expected that the large-system limit for beads-spring polymers starts for comparably smaller system sizes. Still, both cases show that a definite answer is not possible: both fit attempts yield qualitatively good but contradicting results. This is expected from the discussion of particle condensation in Sec. 2.3.1 and suggests that the available data is still within an intermediate regime. If the expected large-system limit is recovered for flexible polymer aggregation, a clear

Table 3.3.: Fit results for the finite-size scaling of $N = 13$ lattice and bead-spring polymer aggregation with monomer density $\rho_m = 0.01$.

Model	$T_0 + aM^{-1/3}$			$T_0 + aM^{-1/4}$		
	Range	T_0	Q	Range	T_0	Q
Lattice	[135: 524]	1.8655(5)	0.98	[242: 524]	1.920(2)	0.66
Bead-Spring	[14: 64]	1.606(2)	0.67	[48: 64]	1.63(2)	0.38

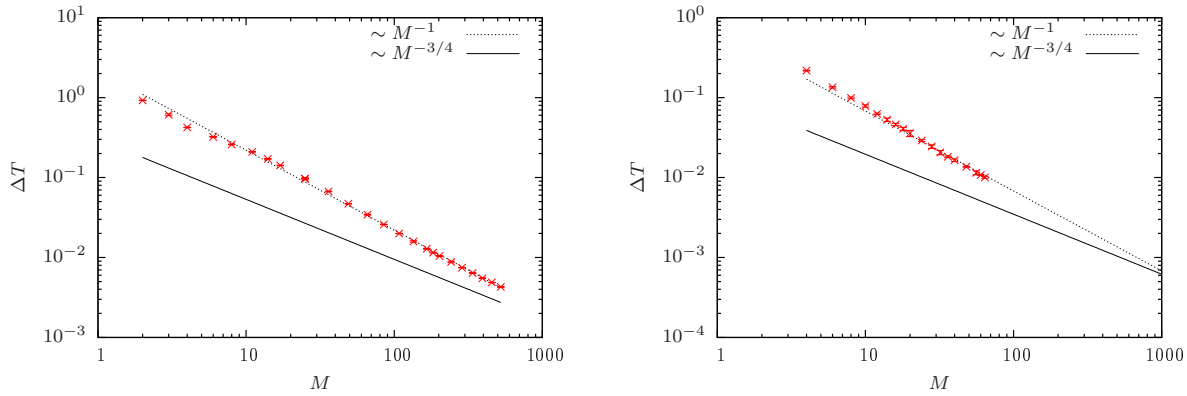

Figure 3.16.: Finite-size rounding of the specific heat peak around the aggregation temperature for $N = 13$ flexible lattice (left) and bead-spring (right) polymers with $\rho_\mathrm{m} = 0.01$. The accessible data in both cases shows a M^{-1} scaling behavior.

signal would require larger system sizes than those currently available.

Similarly, the finite-size rounding of the transition should show different scaling regimes. As rounding of the transition, I consider the half-width of the specific heat peak analogue to Sec. 2.3.2. Since the peak width is supposed to tend towards zero for infinitely large system sizes, one may express the rounding directly proportional to a power-law

$$\Delta T \sim M^{-3p}, \tag{3.11}$$

where again $p = \left\{\frac{1}{3}, \frac{1}{4}\right\}$ for the effective intermediate and large-system regime, respectively. This may be plotted in a double-logarithmic scale showing the power-law exponents as slopes of linear lines. For the considered lattice and bead-spring examples this is shown in Fig. 3.16. In principle, one would expect an intermediate scaling as M^{-1} (dashed line) and a large-system scaling as $M^{-3/4}$ (solid line). In accordance with the discussion of the transition temperature, only an intermediate regime is visible. This regime was also apparent for particle condensation, see Sec. 2.3.2, and was identified as an effective scaling regime with a crossover to the expected large-system scaling regime. Again, larger system sizes are necessary to confirm a condensation-like large-system scaling behavior. If polymers may be considered as extended objects, this would suggest that at least the same number of objects is necessary to reach the large-system limit as for "point-like" objects. Even worse, one would expect that the number of necessary polymers grows with N. The similarities to the case of mesoscopic particle condensation, however, suggest that one may expect the same qualitative behavior for sufficiently large system sizes.

3.4.2. Aggregate size at transition

One remaining canonical observable considered for particle condensation is the scaling of the droplet, or cluster, size in the mixed phase at coexistence. A suitable representation allowed to see (or anticipate) the expected large-system scaling for comparably small system sizes. The cluster size in the mixed phase at coexistence may be estimated from the distribution $P(\eta)$ of cluster sizes, or monomer fractions $\eta = (NM)_D/NM$ in the largest cluster (including $(NM)_D$ monomers) shown in Fig. 3.17 (left). The two peaks in the exemplary distributions

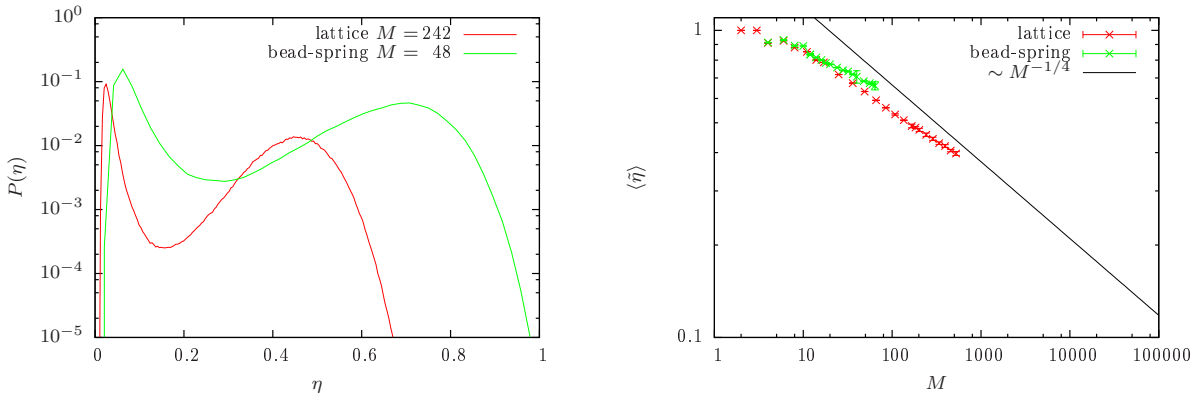

Figure 3.17.: Finite-size scaling of the cluster size at aggregation: exemplary distributions of monomer fraction η in the largest cluster at T_{agg} (left) and scaling of the fraction $\langle\tilde{\eta}\rangle_{T_{\mathrm{agg}}}$ in the transition aggregate (right) for $N = 13$ polymers with $\rho_{\mathrm{m}} = 0.01$.

correspond to the homogeneous, separated phase (small η) and a single macroscopic aggregate with surrounding separated polymers (large η). The non-zero fraction of monomers in the homogeneous phase is partly a consequence of the definition using the largest cluster. This always includes at least N monomers such that $\langle\eta_{\mathrm{hom}}\rangle \to 0$ but is non-zero for finite systems. The local minimum, corresponding to the probability of transition aggregates, decreases with increasing system size in correspondence with the assumptions made in Chap. 2. In addition, these examples already show that one expects with increasing system size smaller average cluster sizes also in the aggregated phase. From the similarities to particle condensation, one may again expect that the large-system transition-aggregate includes $\sim M^{3/4}$ polymers. Consequently, the monomer fraction η should scale as $M^{-1/4}$. Choosing as finite-size transition temperature T_{agg} the peak in the specific heat, the corresponding phases contribute to the distribution with equal weight. Then, the average cluster size in the mixed phase at coexistence may be estimated according to Eq. (2.50): $\langle\tilde{\eta}\rangle_{T_{\mathrm{agg}}} = \frac{1}{\tilde{Z}} \int_{\eta_{\mathrm{min}}} \eta P(\eta)|_{T_{\mathrm{agg}}}\, d\eta$, where η_{min} is defined such that $\tilde{Z} = \int_{\eta_{\mathrm{min}}} P(\eta) d\eta = 0.5$, when the full integral yields 1. The results for both lattice and bead-spring polymers are shown in Fig. 3.17 (right). For small system sizes, where the discrete probability distributions may not even show clear double peaks, this estimate is not very reliable. However, the qualitative conclusion that up to all polymers contribute to the aggregate remains valid. This would correspond to a horizontal line. On the other hand, with increasing number of polymers the $M^{-1/4}$ scaling behavior of the aggregate size (solid line) may be anticipated. This supports the hypothesis of a condensation-like first-order phase transition from a homogeneous, separated polymer phase to a mixed phase of a macroscopic aggregate in equilibrium with surrounding polymer solution.

3.4.3. Microcanonical analysis

A complementary analysis in the microcanonical ensemble [49, 50] promises to contribute to the understanding of dilute polymer aggregation. For a short introduction and application to particle condensation see Sec. 2.5. Figure 3.18 shows estimates of the microcanonical inverse temperature β and its derivative γ for increasing system sizes of lattice polymers ($N = 13$) at constant monomer density $\rho_{\mathrm{m}} = 0.01$. Due to discrete energy states, the application of

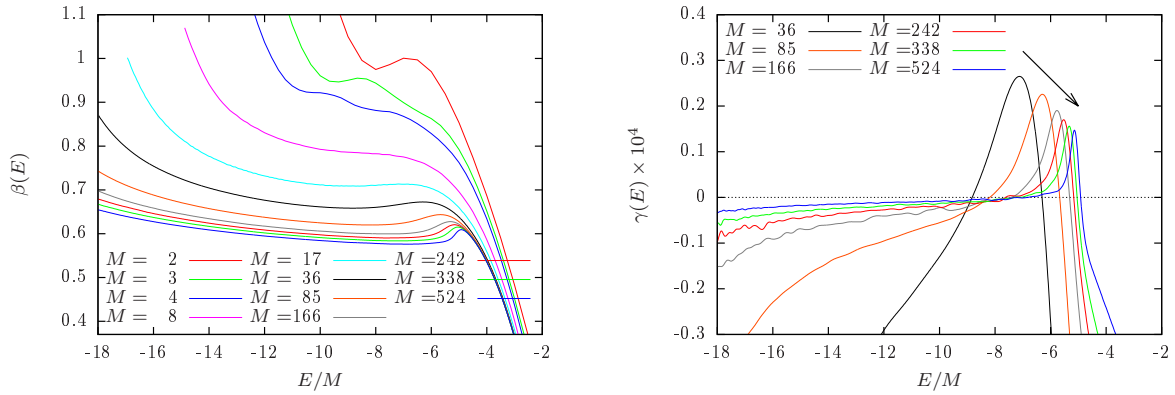

Figure 3.18.: Estimates of the microcanonical inverse temperature β (left) and its derivative γ (right) for lattice polymer systems with $N = 13$ and $\rho_m = 0.01$. The inverse temperature shows signals of nucleation for small M and condensation-like behavior for large M. In the latter case a first-order phase transition is supported by a positive peak in γ approaching zero with increasing system size.

Bézier curves (see Sec. 5.4.1) is limited to large system sizes $M \geq 36$. Interestingly, the inverse temperature shows the characteristic backbending effect for very small ($M = \{2, 3, 4\}$) and quite large system sizes, but not in between. The small system sizes show similarities to the previously observed nucleation hierarchies [53, 64] with oscillations on the curves. For 2 polymers, there is one prominent phase separation process. For 3 and 4 polymers one may anticipate 2 and 3 oscillations, of which only some show a backbending and the corresponding entropic barrier. The vanishing backbending effect, which occurs for the following system sizes, seems to be a lattice effect. For bead-spring polymers there is a small backbending visible for every system size. On the other hand, the bead-spring polymers did not show prominent nucleation hierarchies. This may be explained by the vibrational modes of the harmonic bonds that allow to compensate energy and entropy differences between the sub-aggregated states. For both lattice and bead-spring polymers, the backbending effect becomes prominent again for large system sizes, $M \geq 85$ and $M \geq 32$, respectively. Moreover, the microcanonical signature becomes a first-order signature comparable to the case of condensation (Sec. 2.5). In particular, the derivative γ of the inverse temperature in Fig. 3.18 (right) shows a positive maximum for the aggregation transition which already allows for a first-order classification of the structural transition [49, 55]. Additionally, the decreasing peak height with increasing number of polymers, tending towards zero for $M \to \infty$, allows to speak of a first-order phase transition in the usual sense. This is similarly also observable for bead-spring polymers ($N = 13$) at $\rho_m = 0.01$ already for comparably small system sizes $M \geq 12$ (see also Fig. 4.14).

3.4.4. Crossover from condensation to aggregation

The presented results so far support that the polymer aggregation transition shows strong similarities to particle condensation: a first-order phase transition with coexistence at the aggregation temperature between a homogeneous phase of separated polymers and a mixed phase of a single aggregate in equilibrium with additional separated polymers. The first-order nature of the transition is supported by the microcanonical finite-size scaling but the expected large-

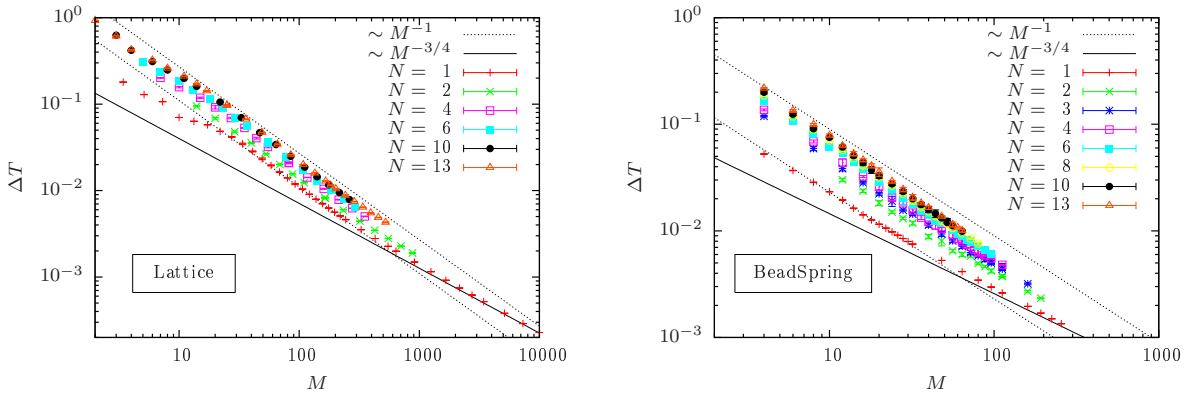

Figure 3.19.: Finite-size rounding from the half-width of the specific heat peak for polymer aggregation of lattice (left) and bead-spring (right) polymers of different length N. The case $N = 1$ corresponds to particle condensation. In all cases an intermediate regime (dashed slope) is clearly visible while the large-system condensation regime may be merely anticipated for $N \neq 1$.

system finite-size scaling behavior is merely visible in the cluster size at the transition temperature. In all other canonical observables, another scaling regime was apparent, consistent with an intermediate scaling regime known from particle condensation. The finite-size rounding of the specific heat peak seemed to best resolve the appearing scaling regimes. Therefore, one may expect a qualitative support of the condensation-aggregation equivalence hypothesis by a visual crossover with increasing polymer length from particle condensation ($N = 1$) to polymer aggregation ($N > 2$). Figure 3.19 shows the finite-size rounding for both polymer models with an increasing polymer length N in a double-logarithmic plot. Different power-law scaling exponents result in different linear slopes. The intermediate scaling regime (dashed slope with M^{-1}) is clearly visible for all polymer lengths. Already the onset of this regime grows with increasing polymer length. From the available data it is suggested that also the onset of the large-system regime grows with increasing polymer length. This explains the difficulties with the numerical results, because the computational complexity increases with longer polymer chains, leading to a smaller maximal number of polymers M. The demand of larger M for the large-system limit is then working in opposite direction, keeping in mind that the necessary system sizes were just about reached for particle condensation. Still, the qualitative crossover from single particle condensation to finite-length polymer aggregation together with the similarities in the intermediate scaling regime support the condensation-aggregation equivalence hypothesis. This would, however, require length scales much large than the typical extension of a single polymer, i.e. $L \gg Nr_0$.

The proposed condensation-aggregation equivalence is limited to the phase separation transition, i.e., the competition between energy minimization and entropy maximization, where microscopic details may be neglected. However, with increasing polymer length N, there arise intra-polymer mechanisms such as the single-polymer collapse. The resulting interplay between collapse and aggregation may be hidden in the considered finite-size scaling analyses but leads to structural changes at the aggregation transition. This microscopic detail may still play a role from a polymer perspective and will be discussed below.

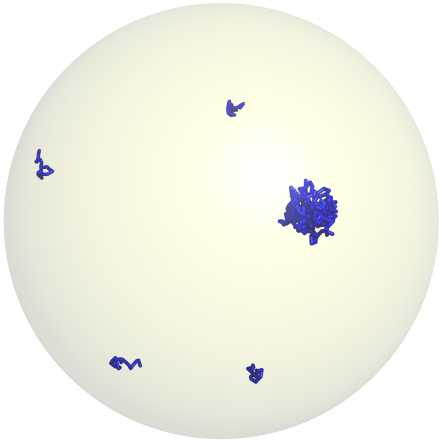

Figure 3.20.: Snapshot from a multicanonical production run of $M = 20$ bead-spring polymers of length $N = 20$ confined in a steric sphere of size $R_S = 30$ [42].

3.5. Density dependence of dilute polymer aggregation in spherical confinement

In order to investigate the effect of density on polymer aggregation in more detail, I consider the model system of flexible bead-spring polymers (Sec. 3.1.2) inside a steric spherical confinement with radius R_S, see Fig. 3.20. The choice may be considered as a safe basis for the study of density effects in finite system. This is because, in contrast to periodic boundary conditions, it allows to systematically vary the density without the possibility that the aggregate or a single polymer may self-interact across the boundaries. On the other hand, the steric confinement (monomers may not cross the bounding shell) brings at most an effective repulsive interaction. For the numerical data, I employ again parallel multicanonical simulations, covering the full energy range of interest in a single simulation. Moreover, this allows to directly sample the entropy of the system, which allows a qualitative insight into the derived arguments. Focusing on the aggregation transition, I refine the energy range using numerical data from dense simulations with short initial parallel tempering simulations. The Monte Carlo moves applied in this study include single-bead displacement, bond rotation, polymer translation, as well as inter- and intra-polymer rearrangement (double-bridging) moves with energy-dependent update ranges (see Sec. 5.6) where possible. The fast conformational entropy increase in larger spheres is partially accounted for by choosing the maximal translation step proportional to the radius of the sphere.

In the following, I first discuss the competition between single-chain collapse and multi-chain aggregation on an example system. Afterwards, I present entropic arguments in the microcanonical ensemble in order to derive a leading-order dependence of the inverse transition temperature on polymer density. Combined with scaling arguments for the finite-size regime this leads to a data collapse of the density dependence for all system sizes with fixed polymer length. Most parts of this section have been published in Refs. [42, 65].

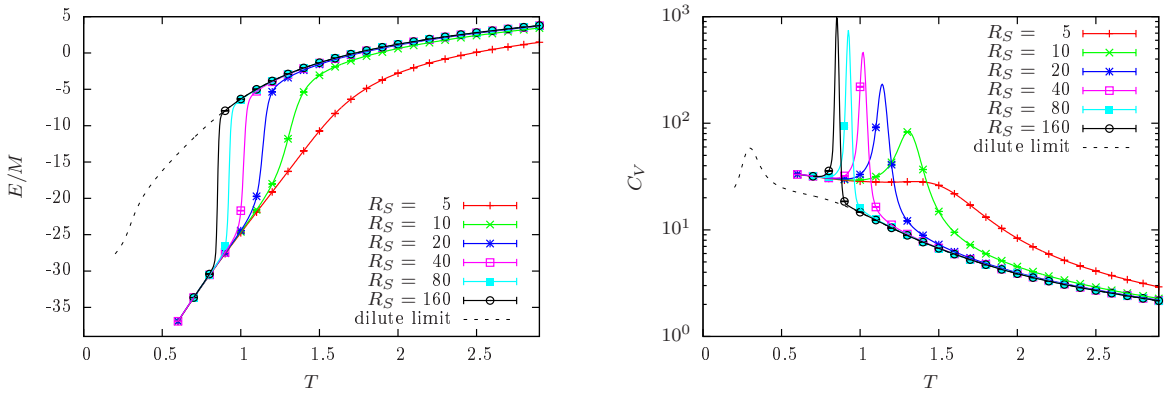

Figure 3.21.: Canonical estimates of the average energy (left) and specific heat (right) for $M = 8$ flexible polymers of length $N = 13$ in spherical confinement. With increasing radius R_S, the density decreases and the aggregation transition shifts to lower temperatures. The dilute limit refers to a single, isolated polymer.

3.5.1. Competition between collapse and aggregation

As discussed in Sec. 3.3, I consider the polymer length N as a system property which is fixed. Above the aggregation transition, the polymers are in a homogeneous phase with independent polymers. In this case, it may be expected that they explore the conformation space independently, behaving like an isolated, single polymer which I refer to here as the *dilute limit*. Moreover, for a fixed number of polymers, an increase in the radius of the confining sphere decreases the density and eventually leads to the dilute limit over the full temperature range for $R_S \to \infty$. For a finite volume, however, there exists an aggregation temperature below which the polymers are no longer homogeneously distributed and independent. In equilibrium, homogeneous aggregation leads to a single macroscopic aggregate possibly with additional independent polymers, see Sec. 3.4. This already shows, that the aggregation temperature has to depend on the density and that in the limit of $\rho \to 0$ it must hold $T_{\text{agg}} \to 0$. In order to illustrate this, consider a system with $M = 8$ homopolymers of size $N = 13$ in different spherical confinements of size R_S. Comparing to the dilute limit, or a single isolated polymer of the same size, it is convenient to consider average single-polymer observables. One suitable choice is the average energy per polymer E/M and the corresponding specific heat per polymer $C_V = \beta^2 \left(\langle E^2 \rangle - \langle E \rangle^2 \right) / M$, where again $\beta = 1/T$ (in units where $k_B = 1$). Another insightful observable is the average end-to-end distance per polymer. This is the average sum over end-to-end distances between the first and the last monomer of each polymer $R_{\text{EE},i}$:

$$\overline{R}_{\text{EE}} = \frac{1}{M} \sum_{i=1}^{M} R_{\text{EE},i}. \tag{3.12}$$

Estimates of the average energy and average end-to-end distance per polymer together with their thermal derivatives are shown in Fig. 3.21 and Fig. 3.22, respectively. The lines are obtained from the raw data of the multicanonical production run by histogram reweighting with a fine temperature resolution. In equidistant temperature steps data points are presented from time-series reweighting with error bars using jackknife error analysis. For details see

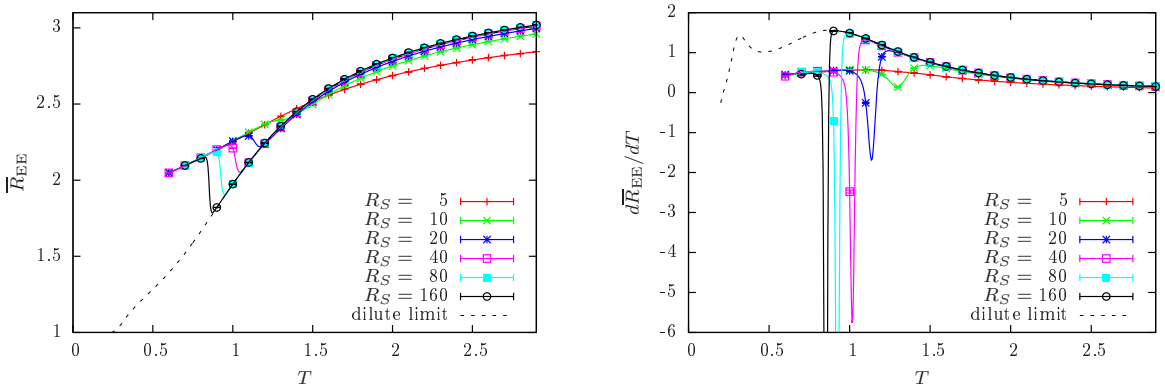

Figure 3.22.: Same as Fig. 3.21 for the average end-to-end distance per polymer (left) and its temperature derivative (right).

Sec. 5.8. The curve "dilute limit" is obtained from a separate multicanonical simulation of a single polymer of length $N = 13$. The aggregation transition temperature may be identified as the temperature where the energy and average end-to-end vector have a sharp jump and the corresponding thermal derivatives show a large, narrow peak. One important observation is the decreasing aggregation temperature with increasing radius, verifying the dilute limit for vanishing density. As expected, above the aggregation temperature almost all curves follow the behavior of an isolated polymer in the dilute limit. This holds true almost directly until the point of equilibrium aggregation. The canonical observables again coincide below the aggregation transition, because then the spherical confinement has almost no effect on the structural properties of the aggregate as long as the confinement is larger than the aggregate.

An exception for the high-temperature, isolated chain regime are rather dense systems (small R_S). In this case, no pronounced aggregation transition occurs and it may be followed that the spherical confinement is of the order of the aggregate and does not allow for separate, independent polymers. This can be understood considering the overlap threshold Φ^* of polymer solutions [3]. For the applied polymer model with average bond length r_0, the volume fraction of a multi-polymer systems is given by

$$\Phi = NM \left(\frac{r_0}{2}\right)^3 / R_S^3. \tag{3.13}$$

On the other hand, the intrinsic volume fraction of a single random coil may be defined as

$$\Phi^* \simeq \frac{N\left(\frac{r_0}{2}\right)^3}{R_{\mathrm{EE}}^3} \simeq N^{1-3\nu} \approx N^{-0.76}, \tag{3.14}$$

where $R_{\mathrm{EE}} \simeq r_0 N^\nu$ is the end-to-end distance of a self-avoiding walk with $\nu \approx 0.588$. A polymer system may be considered dilute if the volume fraction of a multi-polymer system Eq. (3.13) is much smaller than the intrinsic volume fraction of a single random coil Eq. (3.14). A volume fraction of the order of the single Gaussian-coil threshold, $\Phi = \Phi^*$, describes the crossover to the semi-dilute regime, where polymers may no longer be considered independent.

Solving this for the radius of the spherical confinement yields a lower bound for a dilute system:

$$R_S^c \simeq r_0 M^{1/3} N^{\nu}. \tag{3.15}$$

Consequently, if $R_S > R_c$ a multi-polymer system may be considered dilute and around $R_S \approx R_S^c$ a crossover to the semi-dilute regime occurs. For the considered example $R_S^c \approx 6.3$ ($N = 13$, $M = 8$). This is consistent with the deviations observed in Fig. 3.21 and Fig. 3.22 for $R_S = 5$ and to a lesser extent also for $R_S = 10$.

Thus, only for dilute multi-polymer systems the canonical average observables follow the isolated polymer behavior above the aggregation transition temperature. Moreover, Fig. 3.21 and Fig. 3.22 suggest that the individual flexible polymers each follow the collapse transition of the dilute limit down to the point where aggregation suddenly sets in and becomes the dominant physical process. This can be seen best in the average end-to-end distance $\overline{R}_{\mathrm{EE}}$. With decreasing temperature, $\overline{R}_{\mathrm{EE}}$ decreases as expected for polymer collapse with a broad peak around $T \sim 0.9$ in its thermal derivative. Directly at the aggregation transition, however, $\overline{R}_{\mathrm{EE}}$ shows a sharp increase again. This signals that the polymers de-collapse at the aggregation transition in order to form energetically more favorable aggregate structures. Not distinguishing between inter- and intra-chain contacts, the single macroscopic aggregate will assume a spherical shape forming highly entangled, amorphous aggregates (see also Sec. 4.2) rather than patching collapsed polymers together. As a consequence, the average end-to-end distance increases.

The competition between single polymer collapse and multi-polymer aggregation has been observed before. For systems with equal inter- and intra-polymer interactions it has been noticed that they are separate processes but that aggregation dominates especially in the limit of increasing chain length [53]. The dominance of the aggregation transition can be understood by its discontinuous nature opposed to the continuous collapse transition. This discontinuous nature follows from a strong structural variation and can be seen as a sharp jump of, e.g., the end-to-end distance. This is consistent with the presented data, showing furthermore that with varying density the process of single-polymer collapse continues until the density-dependent aggregation temperature. Moreover, Fig. 3.22 shows that the single-polymer collapse is "reverted" at the aggregation transition, unfolding the polymers in order to form equilibrium aggregates with entangled, more extended polymers.

3.5.2. Aggregation transition in the microcanonical ensemble

As mentioned above, the spherical confinement provides a safe and controllable base to study the effect of density on the aggregation transition. In principle, a periodic box would also allow to study this effect. This is discussed for particle condensation in Sec. 2.6. In analogy, here the number of monomers per polymer N is considered as a system property that describes the extension of a polymeric (extended) object. This assumption is useful when comparing conformational entropies with varying density towards the dilute limit. In order to show the generality of the results, we considered three polymer lengths $N = \{13, 20, 27\}$, which show the same qualitative behavior.

The microcanonical ensemble allows to intuitively estimate a leading-order estimate of the relation between aggregation temperature and density. Here, the microcanonical entropy is defined by the logarithm of the total number Ω of configurations with a given (potential) energy: $S(E) = \ln \Omega(E)$. An estimate for $\overline{\Omega}$ may be directly obtained from the parallel multicanonical

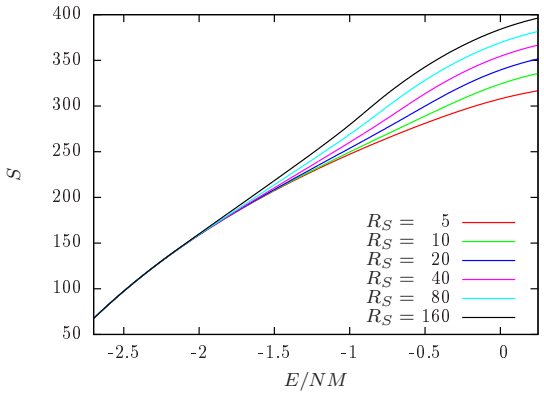

Figure 3.23.: Microcanonical entropy S (up to an additive constant) as the logarithmic number of states $\ln \overline{\Omega}$ for $M = 8$ polymers ($N = 13$), obtained from a microcanonical analysis of the multicanonical data.

simulation in the course of a microcanonical analysis, see also Sec. 2.5 for more details. An example of microcanonical entropies is shown in Fig. 3.23 for the 8×13 system confined in steric spheres with radii R_S. The microcanonical inverse aggregation transition temperature may be estimated as the slope of the hull connecting the aggregated and separated phase [42, 49–51], $\beta_{\text{agg}} = \Delta S / \Delta E$ with $\Delta S = S(E_{\text{sep}}) - S(E_{\text{agg}})$. The arguments are now the same as for the particle gas in Sec. 2.6: consider a fixed number of particles M and varying volume $V \propto R_S^3$. Analogue to an ideal gas, the number of states in the separated phase may be approximated proportional to V^M, i.e.,

$$S(E_{\text{sep}}) \sim \ln \left[\left(\frac{4\pi}{3} R_S^3 \right)^M \right] \propto M \ln R_S. \tag{3.16}$$

The number of states in the aggregated phase will barely be influenced by the confinement, compared to the homogenous separated phase. This can be seen in Fig. 3.23, where $S(E_{\text{sep}})$ dominates over $S(E_{\text{agg}})$ such that $\Delta S \approx S(E_{\text{sep}})$ up to a constant. Moreover, for a fixed number of particles, the energy difference may be considered constant. Then, the inverse microcanonical aggregation temperature depends on the logarithm of the polymer density $\rho = M/V$ [42, 65]

$$T_{\text{agg}}^{-1}(R_S) \sim \frac{S(E_{\text{sep}})}{\Delta E} \sim -\ln \rho + \text{const.} \tag{3.17}$$

This is completely analogue to the results for particle condensation in Sec. 2.6, suggesting that in fact dilute, flexible polymers may be considered as "extended objects". Also, this has been observed recently for two lattice proteins [60] and polymer adsorption [54]. This dependence may be expected to also hold for the aggregation temperature obtained from the peak of the specific heat. This is shown in Fig. 3.24 for $N = 13$ and $N = 20$ and several M with varying density. Each solid line shows a fit to a fixed (N, M) tuple of the form Eq. (3.17). One can see that the derived relation, based on entropic ideal gas assumptions in the microcanonical ensemble, describes the qualitative behavior of the "canonical" aggregation temperature with density. With increasing number of polymers, the slopes of the fits become more similar. The case $N = 27$ is not shown explicitly here but looks similar and will be part of the next

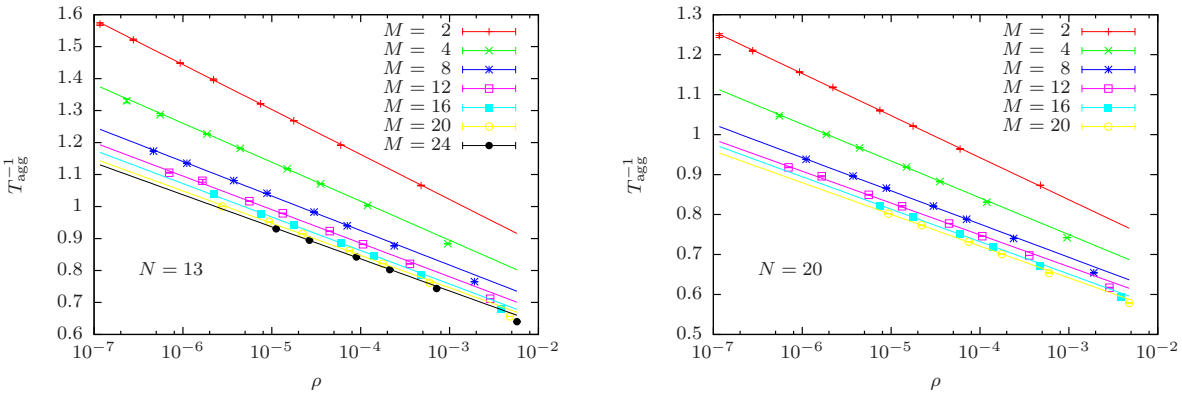

Figure 3.24.: Inverse aggregation temperature as a function of polymer density for several number of polymers M of length $N = 13$ (left) and $N = 20$ (right). The fits support relation Eq. (3.17).

subsection. Tables with aggregation temperatures and the fit results are listed in Ref. [42].

3.5.3. Combining density dependence and finite-size effects

As mentioned before, flexible polymers form an amorphous aggregate with spherical shape. For intermediate polymer numbers, a large fraction of polymers will be included in the aggregate and one may argue that the aggregate should behave similar to a single polymer of length NM below the aggregation transition. Then, the linear extension of the aggregate should scale as $R \sim (NM)^{1/3}$ in this regime of "a few" polymers (see also Sec. 3.4). One measure for this is the total squared radius of gyration

$$R_{\mathrm{gyr}}^2 = \frac{1}{NM} \sum (\boldsymbol{r}_i - \boldsymbol{r}_{\mathrm{cm}})^2 \,, \tag{3.18}$$

where $\boldsymbol{r}_{\mathrm{cm}}$ is the center of mass vector with respect to all monomers. This may be tested for small spheres ($R_S = 30$) where the polymers are dilute enough to form stable aggregates at sufficiently high temperatures. The choice of a small radius is justified considering that the canonical and microcanonical results from the previous discussion in this section show a consistent low-temperature behavior for different radii. Figure 3.25 (left) shows the squared radius of gyration versus the expected scaling function of the total number of monomers $f(NM) = (NM)^{2/3}$. Since this was not the main focus of the study, the multicanonical simulations were restricted to a reduced energy range bounded below by sufficiently small energies for the considered system sizes. One drawback is, that this does not allow to reweight to arbitrarily small temperatures below the aggregation transition. For the comparison of the radius of gyration, a fixed temperature below the aggregation transition is required, which is still within the sampled range of the simulation. As both boundaries vary with the length and number of polymers, this leads to a relatively small sample size in Fig. 3.25 (left): the chosen temperature $T = 0.7$ is too large for the smallest systems to be in the aggregated state but too small for the large systems to be included in the sampled range. Nonetheless, the qualitative data collapse shows that the expectation of the similarity of many- and single-polymer low-temperature extension is valid for finite polymer systems.

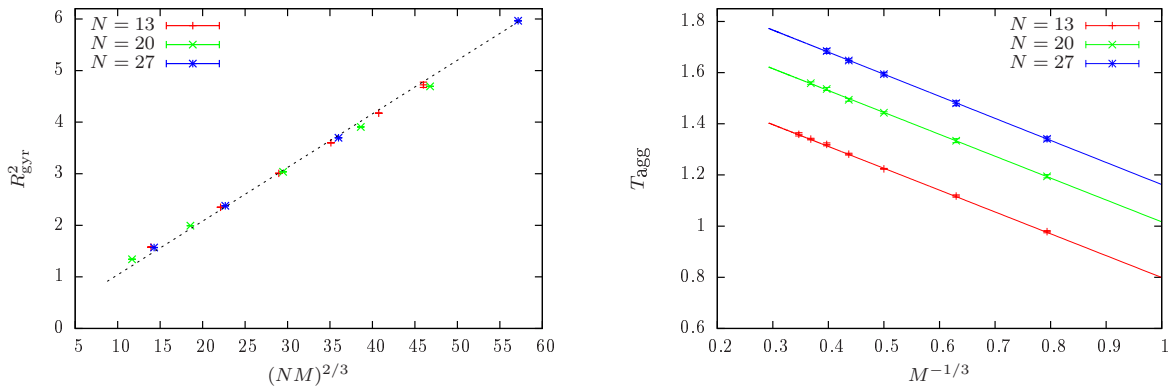

Figure 3.25.: (left) Data collapse of the squared radius of gyration as a function of $(NM)^{2/3}$. The polymer systems are in spheres of size $R_S = 30$ with $N = \{13, 20, 27\}$ at $T = 0.7$. (right) Finite-size effects of aggregation temperature for intermediate M at fixed polymer density $\rho = 10^{-3}$. Error bars are obtained by error propagation and neglect possible systematic deviations due to higher-order corrections.

In order to describe the finite-size dependence of the transition temperature for different M at fixed density, one may thus assume for intermediate-sized systems that a majority of constituents are contributing to the aggregate (compare also Sec. 2.3.3 and Sec. 3.4.2). For flexible polymers, this aggregate is spherically shaped and the linear extension may be related to the radius of a sphere, i.e., the radius of gyration. According to Sec. 3.4, one would expect for the intermediate regime a scaling of the form

$$T_{\mathrm{agg}}(M, N, \rho) \propto \left(1 + s(N, \rho) M^{-1/3} + \mathcal{O}\left(M^{-2/3}\right)\right), \tag{3.19}$$

where $s(\rho)$ is the amplitude of the leading correction and may depend on the density. Having the functional dependence on the density from Sec. 3.5.2 allows to estimate the aggregation transition temperature for any sufficiently small density in the measurement range. The error may be then estimated by error propagation. The result for a polymer density $\rho = 10^{-3}$ is shown in Fig. 3.25 (right) for the three considered polymer lengths. A clear linear behavior of the scaling form Eq. (3.19) is observable. In fact, $s(N, \rho)$ shows a small density dependence, which is expected to converge for small densities. However, a clear conclusion is difficult on the basis of the present data. Since the effects on $s(N)$ are rather small, it is considered constant for the remaining part. The extrapolation to the infinite system size aggregation transition shows the same $\ln \rho$ dependence, but is not physically justified using the intermediate regime. This is, however, consistent with the results for particle condensation in Sec. 2.6 as expected assuming ideal gas behavior and remembering the finite-size scaling discussion in Sec. 3.4. As mentioned before, the available system sizes for polymer aggregation are restricted to the *intermediate* scaling regime, which seems to be quite stable over a broad range of polymer numbers.

If the finite-size corrections may be assumed to be systematic over a broad range of densities ($s(N) \approx const.$), then a rescaling of the aggregation temperature with these corrections would allow for a reduced description independent of system size. Reconsidering the entropy-related microcanonical density dependence, e.g., in Fig. 3.24, and multiplying Eq. (3.19) to the inverse

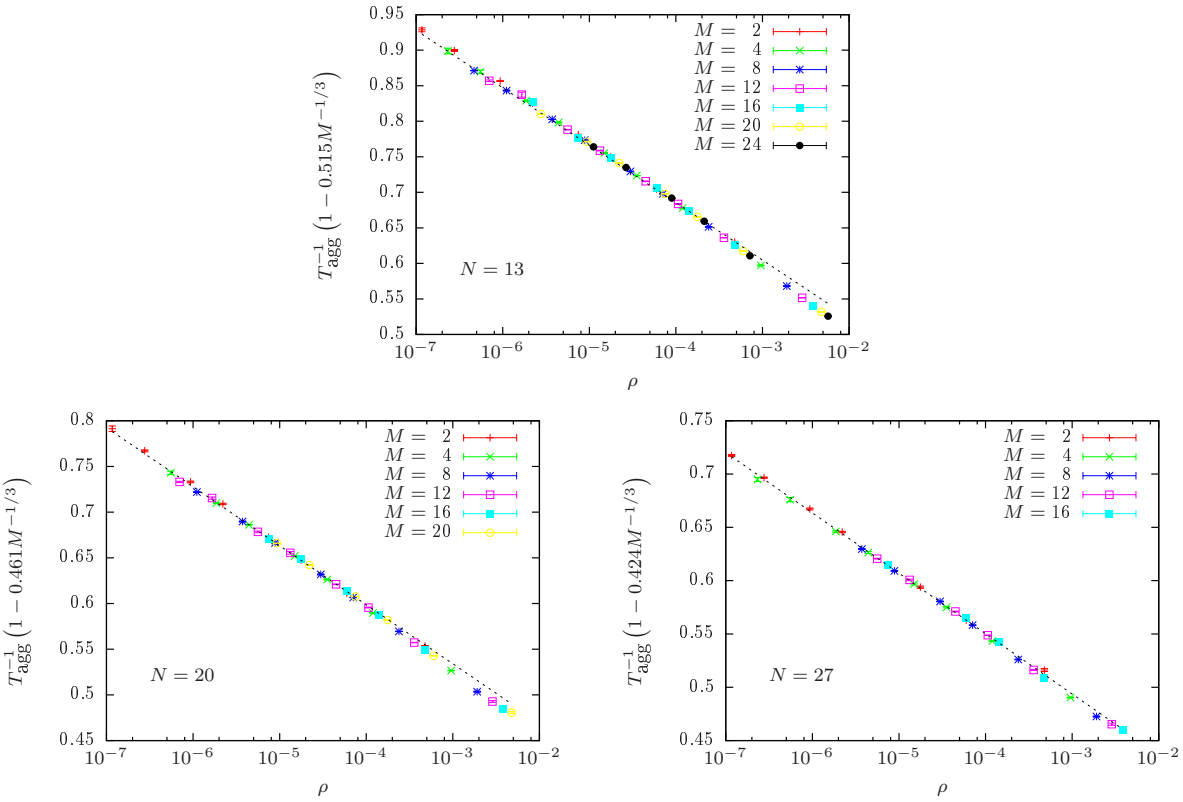

Figure 3.26.: Rescaled inverse aggregation temperature versus density ρ (from Fig. 3.24) for all three polymer lengths $N = \{13, 20, 27\}$.

aggregation temperature reduces the problem to a merely N-depending function of the polymer density. The resulting data collapse is shown in Fig. 3.26 for the three considered cases. Here, the parameter $s(N)$ was assumed constant and its value is denoted on the y-axes of the plots. The dashed line is a guide to the eye. The data collapse on a single line over several orders of magnitude shows that assuming a globular macroscopic aggregate on the one side and an ideal gas approximation on the other side allows for an efficient description of polymer aggregation in the limit of a mesoscopic number of dilute uniform polymers. Moreover, this is consistent with the interpretation of polymers as extended particles for which the same considerations yielded comparable results.

4. The role of stiffness in polymer aggregation

Introducing bending stiffness to (flexible) Θ-polymers results in a model system for an entire class of semiflexible polymers with excluded volume and short-range attraction. The interplay of excluded volume, short-range attraction and stiffness causes a multitude of structural phases. Early studies of the role of stiffness for a single polymer include mean-field calculations [83] and PERM chain-growth simulations of lattice models [84]. Depending on the considered model, this structural phase diagram may become quite extensive – ranging from globular to toroidal structures – as was shown for tubelike polymers [85–89] and bead-spring polymers [90, 91]. For sufficiently large stiffness, the wormlike chain model is recaptured, which only considers stiffness and neglects self-avoidance and self-attraction.

Studying aggregation of coarse-grained semiflexible Θ-polymers allows to unravel which properties may be reproduced already with a simple, generic model relying merely on self- and mutual avoidance, short-range attraction and stiffness. This includes studies of general mechanisms for peptide [51, 52, 92, 93] and homopolymer [42, 53, 65, 94, 95] aggregation. Focusing on the effect of stiffness, it was possible to show that, depending on stiffness, the aggregation transition of homopolymers may be accompanied by an additional freezing transition [53] and that stiffness plays a key role for the structural phases of polymer aggregates [95] (see Sec. 4.2). For sufficiently large stiffness, semiflexible Θ-polymer bundles show consistent results with *interacting* wormlike chains. Examples are the study of unbinding transitions for two and more parallel filaments [76, 96] and the twisting of filaments in a wormlike bundle model [97–99]

Using analytical and numerical techniques enables to understand or extend mechanisms, check claims from experiments, and in the best case make predictions for experiments. For example, structural motifs similar to those of twisted bead-spring or wormlike bundles occur after bundling into mature amyloid fibrils [100] and upon adsorption onto nano-wires [101]. Furthermore, it was argued that the free-energy barrier of amorphous aggregates is lower than for ordered structures, such as amyloid fibrils [102]. This was qualitatively verified comparing amorphous aggregates and polymer bundles of semiflexible bead-spring polymers [95] (see Sec. 4.2 and Sec. 4.3).

This chapter is organized in the following way. The model of a bead-spring Θ-polymer is extended to semiflexible polymers in Sec. 4.1. The resulting structural phases, from amorphous aggregates to polymer bundles, are discussed in Sec. 4.2 for 2-8 (short) polymers. This includes the introduction of an effective bundle parameterization in order to characterize low-temperature conformations. In Sec. 4.3, I investigate the effect of stiffness on the fixed-density finite-size scaling in an accessible (mesoscopic) regime. In extension to the previous discussion, Sec. 4.4 recaptures the effect of density on aggregation of semiflexible polymers.

4.1. Semiflexible polymer model

In order to study the effect of stiffness on polymer aggregation, I restrict the discussion to the case of bead-spring polymers in continuous space. While it is in principle possible to introduce

stiffness also in a lattice model, this would of course introduce additional discretization effects. Semiflexible polymers may be successfully described by the Kratky-Porod or wormlike-chain (WLC) model [103, 104]. This has been applied to the study of structural and dynamic properties of DNA, RNA, microtubules, intermediate filaments and actin filaments, to name a few prominent examples. The WLC model is a continuum formulation of a semiflexible polymer with contour length L, neglecting microscopic details on the monomer scale, and modeling stiffness by an energy attributed to thermally excited bending. If the contour is described by the continuous curve $R(s)$, the curvature-depending Hamiltonian may be written as

$$\mathcal{H}^{\mathrm{WLC}} = \frac{\tilde{\kappa}}{2} \int_0^L ds \left(\frac{\partial^2 R(s)}{\partial s^2} \right)^2 , \tag{4.1}$$

where $\tilde{\kappa}$ is the bending stiffness. Here, one usually introduces three regimes depending on the bending stiffness or, more commonly, the persistence length $l_p = \tilde{\kappa}/k_{\mathrm{B}}T$, which may be associated with the tangent-tangent correlation length along the polymer. Usually, $r_0 \approx l_p \ll L$ refers to flexible polymers, $r_0 \ll l_p \approx L$ refers to semiflexible polymers, and $l_p \gg L$ refers to stiff polymers.

This formulation allows to treat certain problems analytically. The complexity of (bio-) polymer physics is simplified by considering the isolated property of stiffness, neglecting self-avoidance and also self-attraction. This is a valid assumption for many semiflexible polymers and in the limit of stiff polymers. Considering Θ-polymers, however, introduces a short-range self-attraction with the corresponding collapse transition, which may not be easily covered by the WLC model. For several interacting wormlike chains, one usually introduces so-called cross-linkers in order to model bundle transitions or networks.

4.1.1. Extension of the bead-spring model to semiflexible polymers

The success of the wormlike chain shows that the isolated property of bending stiffness alone allows to describe a wide range of physical processes. Additionally considering the properties of self-avoidance and short-range self-attraction yields a semiflexible Θ-polymer model. This may be modeled with coarse-grained monomers connected by bonds and is no longer a continuum formulation but a discrete polymer model in continuous space. Thus, the bending stiffness needs to be formulated in terms of bonded monomers. The bonds may be related to local tangents of the curve and may be written in terms of unit bonds $\boldsymbol{b}_i = b\boldsymbol{u}_i$. For equally long bonds the WLC Hamiltonian may be rewritten as $\mathcal{H} = (\tilde{\kappa}/b) \sum (1 - \boldsymbol{u}_i \boldsymbol{u}_{i+1})$ [105], which is also known as the Heisenberg chain model up to a constant. Rewriting the scalar product in terms of the angle θ_i between neighboring bonds, i.e., $\boldsymbol{u}_i \boldsymbol{u}_{i+1} = \cos\theta_i$, leads to $\mathcal{H} = (\tilde{\kappa}/b) \sum (1 - \cos\theta_i)$. Formally, this only holds for equally long bonds.

For models with elastic, harmonic bonds like the bead-spring polymer one may argue that the average bond length is the relevant length scale and may be associated with the constant bond length. Moreover, it suffices to consider a potential that is motivated by the isolated properties of the wormlike chain, which itself remains a model. Then, the bead-spring model from Sec. 3.1.2 may be extended to semiflexible Θ-polymers by introducing a WLC-like bending energy term

$$E_{\mathrm{bend}} = \kappa \sum (1 - \cos\theta_i) , \tag{4.2}$$

where κ is the bending stiffness that may be merely qualitatively related to the bending stiffness $\tilde{\kappa}$ in the WLC formulation Eq. (4.1). Several polymers then naturally interact via inter-polymer

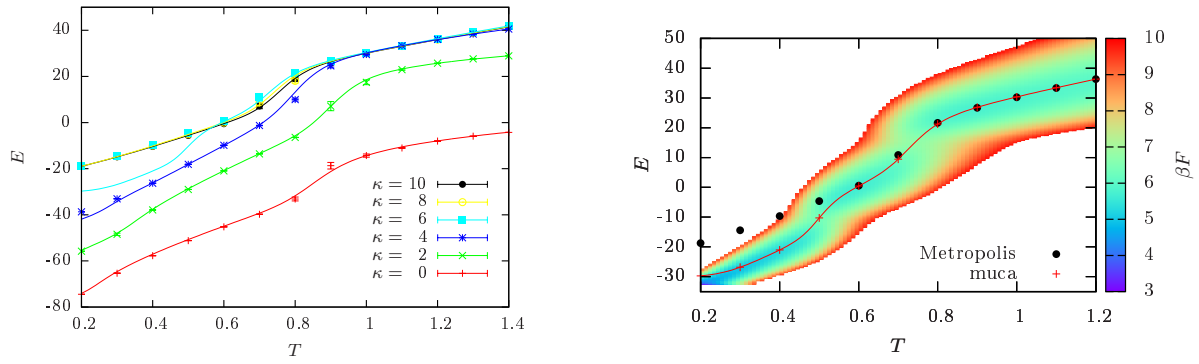

Figure 4.1.: (left) Comparison of independent Metropolis simulations (data points) with canonical expectation values reweighted from parallel multicanonical data (solid lines) for different bending stiffness κ ($M = 2$, $N = 13$, $\rho = 10^{-3}$). (right) The case $\kappa = 6$ in its free-energy landscape and multicanonical data points.

Lennard-Jones contacts (with energy scale ϵ) as for flexible polymer aggregation in Chap. 3. In the following, I will refer to rather flexible polymers for $0 \approx \kappa < N\epsilon$ and to rather stiff polymers for $\kappa \approx N\epsilon$. This classification is in terms of energies and may be related to a persistence length only by considering the temperature scale (see above).

The Monte Carlo updates I applied are the same as for flexible polymers in Sec. 3.1.2. All numerical results in this chapter are again obtained from parallel multicanonical simulations. As usual the data points with error bars are calculated using time-series reweighting with jackknife error analysis and the lines with higher resolution using histogram reweighting. In order to test the implementation of the model, I performed independent Metropolis simulations for 2 polymers of length $N = 13$ for several κ with local updates (single monomer displacement) only. For each temperature, 20 000 measurements were recorded with 52 000 updates in between. Each set of data points (fixed κ) is obtained from a single simulation starting at high temperatures and step by step cooling down until the lowest temperature is reached. Figure 4.1 shows results from both Metropolis and multicanonical simulations. The left figure shows the total (conformational) energy E for several κ values, verifying both implementation and analysis: the Metropolis data points are scattered around the muca estimates from histogram reweighting (solid line). The error bars are obtained including an integrated autocorrelation time analysis. It can be seen that the high-temperature regime shows strong agreement, but differences occur in the low-temperature regime especially for $\kappa = 6$. It may be argued that the Metropolis algorithm has difficulties in the latter temperature range, due to low acceptance rates and large conformational or entropic barriers.

The extreme case $\kappa = 6$ is shown again in Fig. 4.1 (right) together with the microcanonical free-energy landscape. The microcanonical free energy is related to the energy probability distribution up to a constant, $\beta F(E) = -\ln P(E)$, which can be obtained directly from the multicanonical simulation (see also Sec. 2.4). It can be seen that the data points from the multicanonical simulation (red crosses and line) follow the free-energy minimum. On close sight, there are two temperature regions with two local free-energy minima: one around $T \approx 0.73$ and one around $T \approx 0.5$. The larger temperature may be associated to the aggregation temperature and the lower temperature is connected to the formation of entangled hairpins (see discussion

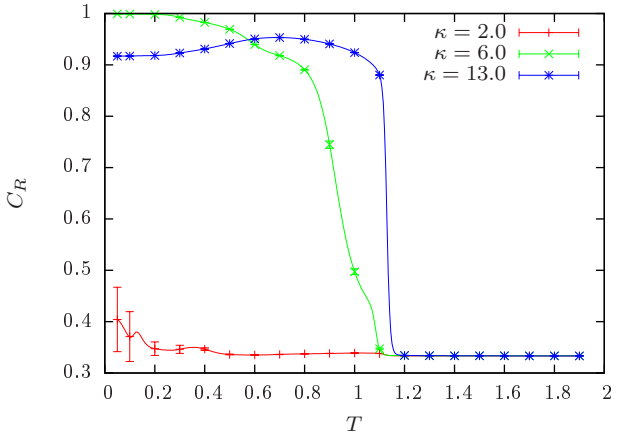

Figure 4.2.: Example of the end-to-end correlation parameter C_R, distinguishing between amorphous aggregates ($\approx 1/3$) and correlated bundles (≈ 1) for 8 semiflexible polymers of length $N = 13$ and $\rho_\mathrm{m} = 10^{-3}$. More flexible polymers ($\kappa = 3$) aggregate into an amorphous phase; rather stiff polymers ($\kappa = 13$) aggregate directly into bundles.

in Sec. 4.2.1). For low temperatures, the Metropolis simulation seems to get stuck in a conformation that may not be easily transformed into the corresponding equilibrium conformation. This is a known problem of local updates in canonical simulations at low temperatures. The overall picture, however, supports that the multicanonical simulation samples the state space fulfilling both ergodicity and detailed balance.

4.1.2. End-to-end correlation parameter

As mentioned in Chap. 3 one observes amorphous aggregates for flexible polymers. For sufficiently stiff polymers, however, one expects to observe polymer bundles. In order to distinguish amorphous aggregates from bundle or fibril structures, we introduce an order parameter [95] that measures the correlation of the normalized end-to-end vector $\boldsymbol{R}_i$ of each polymer via the scalar product:

$$C_R = \frac{2}{M(M-1)} \sum_{i<j} \left(\hat{\boldsymbol{R}}_i \cdot \hat{\boldsymbol{R}}_j \right)^2 . \tag{4.3}$$

If all polymers are separated and freely moving, their relative orientations are independent of each other. Integration over the full angular space of independent vectors results in $C_R = 1/3$ (see the following argumentation). The same holds approximately for the amorphous aggregate, where the polymers may be considered uncorrelated. On the contrary, forming a bundle, i.e., all polymers aligned in a cylindrical shape, results in a strong correlation of the end-to-end vectors. Then, the average angle between $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$ will be almost zero, such that $\cos(\alpha) \approx 0$ and $C_R \to 1$. For an example with 8 polymers of length $N = 13$ see Fig. 4.2. Alternatively, one may consider the nematic order parameter [106] at the additional cost of computing a full bond-bond interaction tensor.

The bounding cases of correlated and uncorrelated vectors may be estimated by taking a look at two unit vectors $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$. Starting with the correlated case, both vectors will be

(anti-) parallel and the enclosed angle α vanished in the limit of completely correlated vectors, such that the average squared scalar product will be 1. The same will be true if the number of vectors is increased and all of them are correlated.

Consider now both unit vectors to be completely independent of each other. The average squared scalar product is evaluated by integration over the full configuration space of both vectors

$$\langle(\cos\alpha)^2\rangle = \int \frac{dV_i dV_j}{16\pi^2}(\boldsymbol{R}_i\boldsymbol{R}_j)^2 = \int d\theta_i d\varphi_i d\theta_j d\varphi_i \frac{\sin\theta_i \sin\theta_j}{16\pi^2}(\cos\alpha)^2\,, \tag{4.4}$$

with a transformation to spherical coordinates $\boldsymbol{R}_i = (\sin\theta_i\cos\varphi_i, \sin\theta_i\sin\varphi_i, \cos\theta_i)$ and the resulting Jacobian $\sin\theta$. Inserting all the definitions into the squared scalar product yields

$$\begin{aligned}
\cos^2\alpha = {} & \sin^2\theta_i \sin^2\varphi_i \sin^2\theta_j \sin^2\varphi_j \\
& + 2\sin^2\theta_i \sin\varphi_i \cos\varphi_i \sin^2\theta_j \sin\varphi_j \cos\varphi_j \\
& + 2\sin\theta_i \cos\theta_i \cos\varphi_i \sin\theta_j \cos\theta_j \cos\varphi_j \\
& + \sin^2\theta_i \sin^2\varphi_i \sin^2\theta_j \sin^2\varphi_j \\
& + 2\sin\theta_i \cos\theta_i \sin\varphi_i \sin\theta_j \cos\theta_j \sin\varphi_j \\
& + \cos^2\theta_i \cos^2\theta_j .
\end{aligned}$$

Since all angles are independent of each other, the integrals may be evaluated independently. For this task, I make use of the following list of known integrals:

$$\begin{aligned}
&\int dx \sin x = -\cos x && \int dx \cos x = \sin x \\
&\int dx \sin^2 x = \frac{1}{2}(x - \sin x\cos x) && \int dx \cos^2 x = \frac{1}{2}(x + \sin x\cos x) \\
&\int dx \sin^3 x = \frac{1}{12}(\cos 3x - 9\cos x) && \int dx \sin x\cos x = -\frac{1}{2}\cos^2 x \\
&\int dx \sin^2 x\cos x = \frac{1}{3}\sin^3 x && \int dx \sin x\cos^2 x = -\frac{1}{3}\cos^3 x
\end{aligned}$$

Integrating over $\theta \in [0,\pi]$ and $\varphi \in [0,2\pi]$ causes the contributions with linear terms of sin or cos to vanish and one remains with

$$\begin{aligned}
\langle(\cos\alpha)^2\rangle &= \int \frac{1}{16\pi^2}\{\sin^3\theta_i \cos^2\varphi_i \sin^3\theta_j \cos^2\varphi_j \\
&\qquad + \sin^3\theta_i \sin^2\varphi_i \sin^3\theta_j \sin^2\varphi_j \\
&\qquad + \sin\theta_i \cos^2\theta_i \sin\theta_j \cos^2\theta_j\} \\
&= \frac{1}{16\pi^2}\left\{\left(\frac{4}{3}\right)^2\pi^2 + \left(\frac{4}{3}\right)^2\pi^2 + \left(\frac{2}{3}\right)^2(2\pi)^2\right\} \\
&= \frac{1}{3}.
\end{aligned}$$

With the proper normalization in Eq. (4.3), the same result is also valid for an increasing number of vectors, if all of them are completely uncorrelated.

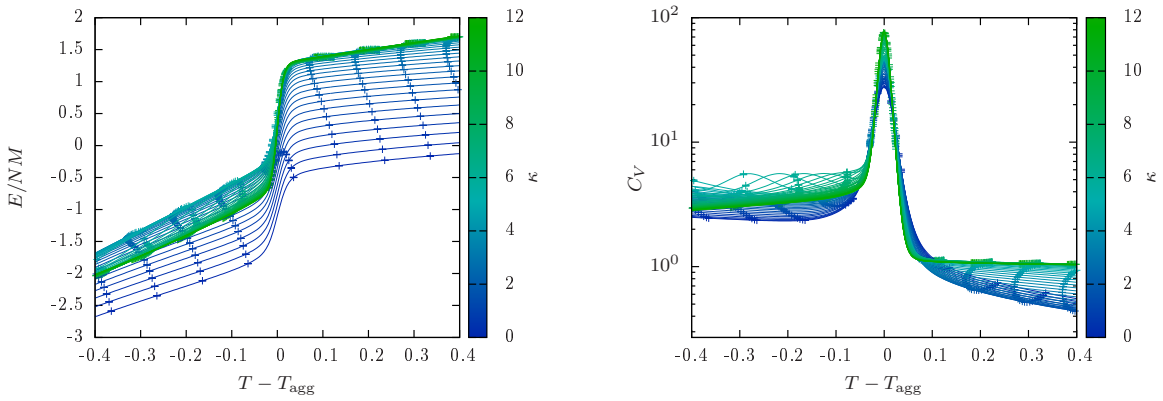

Figure 4.3.: Effect of stiffness on the canonical estimates of the average energy (left) and specific heat (right) for 8 semiflexible polymers of length $N = 13$ ($\rho_\mathrm{m} = 10^{-3}$). The temperature is shifted around the aggregation temperature T_agg and the graph color encodes the bending stiffness κ in the interval $[0, 12]$ with $\Delta\kappa = 0.2$.

4.2. Structural phases of semiflexible polymer aggregates

One key question to be answered is the effect of stiffness on the structure of Θ-polymer aggregates. The arising complex interplay of collapse and aggregation with stiffness leads to generic κ-T structural phase diagrams of an entire class of semiflexible polymers. This section focuses on the discussion of these structural phase diagrams for 2-8 semiflexible polymers of length $N = 13$. Of particular interest will be the arising intermediate-stiffness regime, the occurrence of twisted bundles for more stiff polymers at low temperatures, and a comparison of free-energy barriers for amorphous aggregates and polymer bundles. Most of the results in this section are published in Ref. [95].

In order to relate the aggregation transition of semiflexible Θ-polymers to flexible polymers ($\kappa = 0$, see also Chap. 3), Fig. 4.3 shows canonical estimates of the average energy and the specific heat C_V per monomer for 8 polymers with different κ (encoded in the line color). The temperature axis is shifted to center around T_agg, defined as the temperature of the largest peak in C_V. In all cases, the data display a sharp, discontinuous transition when the polymers aggregate from an entropy dominated, soluble regime at high T into an energy dominated aggregated regime at low T. On close sight, one may observe that the size of the energy jump and the height of the specific heat peak increase with stiffness. Below the aggregation transition, further less pronounced peaks are visible for intermediate κ values, indicating continuous transitions subject to discussion below. The energy clearly shows the involved structural transitions but seems unsuitable to distinguish the possible structures of the aggregate. The same holds for the phase separation parameter Γ^2 (see Sec. 3.1), which only shows small differences, see also Ref. [65]. Therefore, I make use of the end-to-end correlation parameter introduced above in order to distinguish the occurring structural motifs.

4.2.1. Structural phase diagrams for finite semiflexible polymer systems

An overview over the involved structural phases of small semiflexible polymer systems with $M = \{2, 4, 8\}$ polymers of length $N = 13$ is shown as generic κ-T "phase" diagrams in Fig. 4.4. The monomer density is in all cases $\rho_{\mathrm{m}} = 10^{-3}$. Due to the finite systems, the transitions between structural phases should not be confused with thermodynamic phase transitions. As discussed in Sec. 3.4, the thermodynamic limit may be recovered for the aggregation transition in the limit $M \to \infty$ (see also Sec. 4.3). From the discussion of flexible polymer aggregation, one may expect with increasing polymer number a mixed aggregated phase consisting of a single macroscopic aggregate in equilibrium with separated polymers. For the considered system sizes and small density, the aggregated conformations include all polymers. The discussion of the corresponding finite-size structural phase diagram is thus not applicable to the thermodynamic limit but gives insight into the transition mechanisms and structural properties involving few polymers. This may be relevant also for non-equilibrium considerations of semiflexible polymers, where inhomogeneous aggregation results in local interaction of only relatively few polymers. For semiflexible polymers this may lead to network formation and twisted structures. Examples include the finite thickness of DNA bundles observed in experiment and simulation [107], as well as finite thickness of F-actin bundles [108, 109]. The (equilibrium) structural phase diagrams presented here highlight what happens in the aggregation process of a finite number of polymers for an entire class of semiflexible polymers over a wide range of stiffness and temperature.

The diagrams are obtained from parallel multicanonical simulations at fixed κ values in the range $[0, 14]$ with $\Delta\kappa = 0.2$. This includes both flexible ($\kappa = 0$) and rather stiff ($\kappa \approx N\epsilon$) polymers. Exemplary simulations of shorter and longer chain length support the obtained picture. The structural phase diagrams combine the landscape of the end-to-end correlation parameter C_R with transition points obtained from the peaks of the specific heat (black dots) and of the thermal derivative of the phase separation parameter (blue squares). In the latter case, the color encodes the strength of the signal from rather strong signals (dark blue) to rather weak signals (light blue), which in some cases may be merely fluctuations without physical significance. Next to the diagrams, typical conformations are presented from different structural regions for selected κ and T. In principle, I distinguish between the separated or soluble phase (S), in which all polymers are independent of each other; the aggregated phases (A) with polymers located close to each other; and the "frozen" phases (F) as the corresponding low-temperature regions with rather well-ordered structures. The aggregated and frozen phases are further differentiated into amorphous aggregates (A_1, F_1) with $C_R \approx 1/3$ for rather flexible polymers (small κ), and polymer bundles (A_2, F_2) with $C_R \approx 1$ for stiffer polymers. For referencing purposes, the frozen phases are partially further subdivided with a second index.

The amorphous aggregates (A_1) show uncorrelated structures for all three cases, visible in $C_R \approx 1/3$ and looking at the corresponding typical conformations. The compact, spherical structure suggests that close packing is the driving formation mechanism, maximizing monomer-monomer contacts. The line-shape of the polymers seems unimportant. For rather flexible polymers, this seems similar to the collapsed state of a single polymer or a cluster of non-bonded monomers. Decreasing the temperature even further results in additional freezing-like transitions, again similar to the single-polymer case [69, 86–88, 90] but with the additional possibility to wrap around each other (F_1). On the contrary, stiffer polymers form correlated polymer bundles (A_2), aligning elongated polymers in parallel. Here, clearly the linear na-

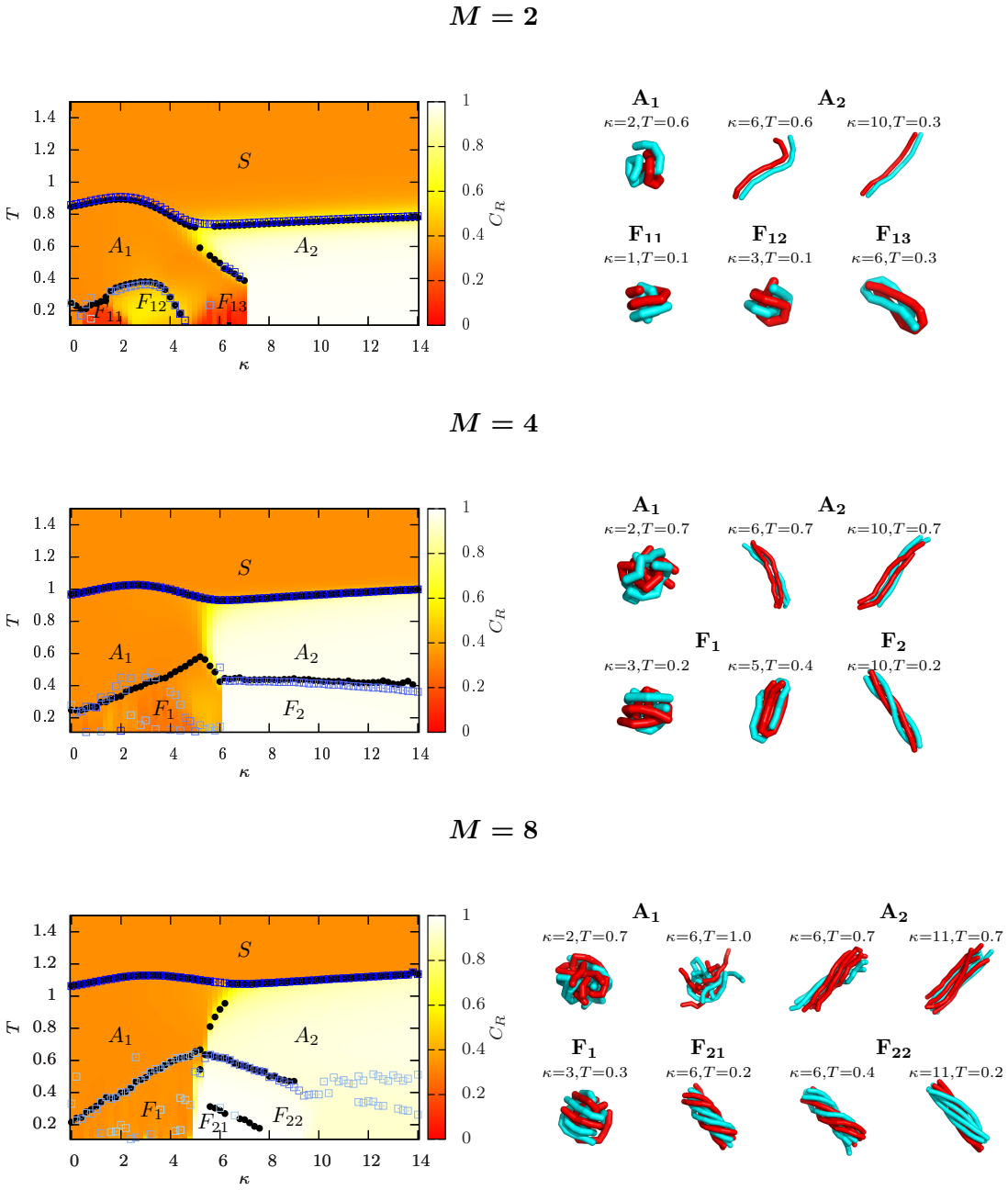

Figure 4.4.: Structural phase diagrams for 2, 4 and 8 polymers with $N = 13$ monomers each ($\rho_m = 10^{-3}$). The surface plot shows the end-to-end correlation parameter C_R; the symbols mark the maxima of the heat capacity (black dots) and of the temperature derivative of the phase separation parameter Γ^2 (blue squares). Several structural phases are identified, namely S (soluble), A (aggregated) and F (frozen), and typical conformations for the low-temperature phases are presented. The figure is partially adapted from Ref. [95].

ture of the polymer is important. Within these large-temperature bundles, the polymers may fluctuate both along the bundle direction and perpendicular. Similar bundle structures were observed in simulations of semi-dilute DNA solutions as explicit semiflexible polymers [107] with comparable parameters to those considered here. For lower temperatures, the bundles tighten until they may be describes as *twisted bundles* (F_2) recapturing the qualitative behavior of the wormlike bundle model [97] for sufficiently large stiffnesses at fixed N. This is already observed for 4 polymers and becomes more prominent for 8 polymers. The twist may be explained by a large energetic gain from maximizing Lennard-Jones contacts compared to the energetic loss from bending (see discussion below in Sec. 4.2.4). This is consistent with analytical considerations of twisted fibers [98, 99] and numerical studies with several tubelike polymers [94]. For rather stiff Θ-polymers, the (discrete) wormlike chain seems to become a good approximation. This allows to relate the persistence length and bending stiffness to leading-order as $l_p/r_0 \approx \kappa/k_BT$. For a large range of the twisted bundle regime in Fig. 4.4, the length ratio l_p/Nr_0 is of order unity. This matches typical scales of amyloid fibrils, where the pitch length is of the order of the protofibril length (which, however, are both much larger) [110]. Additional specific interactions such as hydrogen bonds and hydrophobic effects may stabilize or destabilize structural motifs.

In all three cases, an intermediate stiffness regime is observable roughly between $\kappa \approx 4 \ldots 8$, where both uncorrelated and correlated aggregates are present for a fixed κ. The details, however, differ noticeably from 2 to 8 polymers. For 2 polymers, lowering the temperature in this region leads to an initial correlation into polymer bundles followed by a decorrelation into hairpin-like structures (F_{13}). The entangled hairpins may have a slight twist and in some cases entangle perpendicular ($C_R \approx 0$ for 2 polymers only). A similar picture is obtained for 4 polymers, while the crossover along κ shows that within the narrower regime both amorphous aggregates and polymer bundles are observable. For low temperatures, again hairpin-like structures are achievable. However, the case of 4 polymers already differs from 2 polymers and indicates a qualitative change in the intermediate regime. The case of 8 polymers then demonstrates the reverse situation, where first uncorrelated aggregates A_1 form right below the aggregation transition, followed by a correlation of polymer end-to-end vectors for lower temperatures resulting in polymer bundles A_2. Lowering the temperature even further, the bundles undergo an additional structural transition into the frozen twisted bundle phases. Follow for example the temperature along $\kappa = 6$, which is also supported with typical conformations in Fig. 4.4. Varying κ at fixed low T leads to a narrow crossover in all cases.

4.2.2. Microcanonical analysis in the intermediate stiffness regime

As discussed in Sec. 2.5, a microcanonical analysis [49, 50] may provide additional information for structural transitions and is a suitable tool to classify the order of a structural transitions in finite systems [55]. This is now particularly useful in order to shed some light into the finite-size transitions within the intermediate stiffness regime.

An estimate of the microcanonical entropy $S(E)$ (in terms of the potential energy) may be obtained directly from the multicanonical production run. Using numerical derivatives (see Sec. 5.4.1) yields the microcanonical inverse temperature $\beta(E) = \partial S(E)/\partial E$ and its derivative $\gamma(E) = \partial\beta(E)/\partial E$. This is shown in Fig. 4.5 for the three cases $M = \{2, 4, 8\}$ inside the intermediate regime. In all cases, the aggregation transition at larger energies shows a first-order like signature ($\gamma > 0$ peak). For 2 polymers an additional first-order transition occurs at lower energies. This may be associated to the transition $A_2 \to F_{13}$ from an extended "bundle"

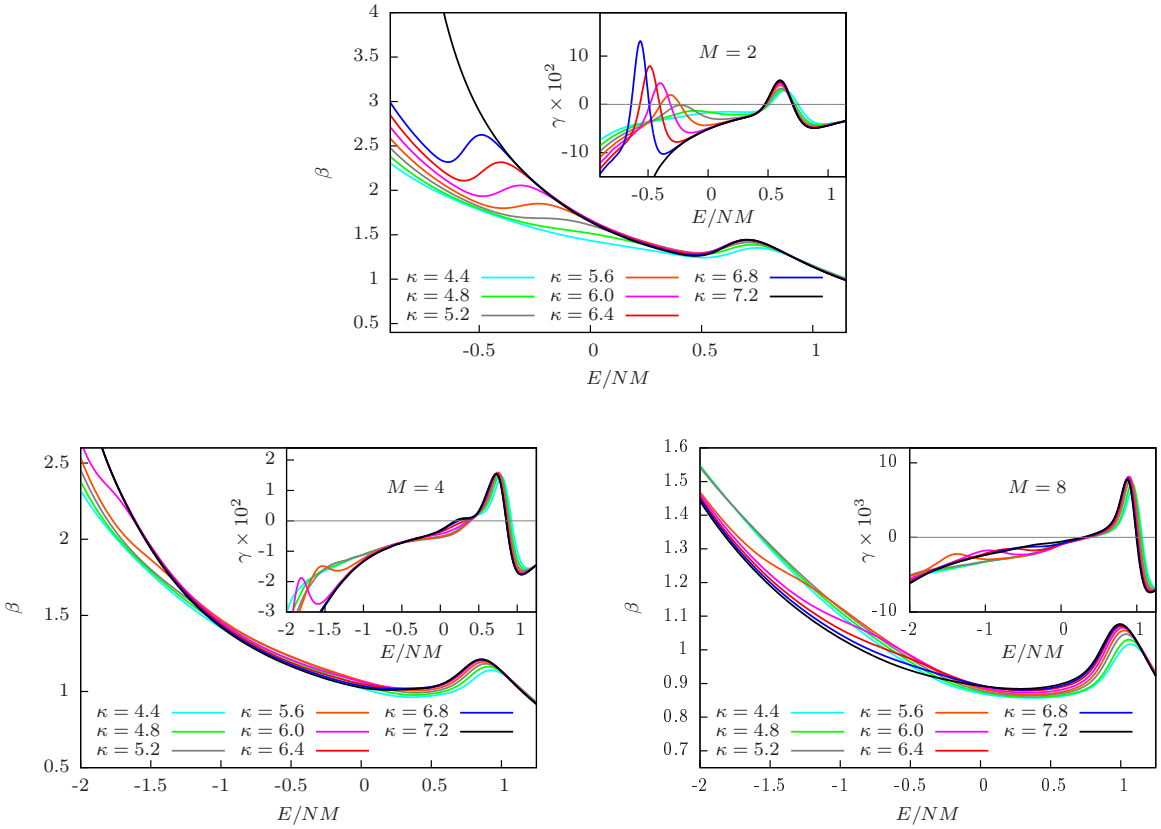

Figure 4.5.: The first (β) and second (γ) derivatives of the microcanonical entropy for $M = 2$, $M = 4$ and $M = 8$ polymers with $N = 13$ show the occurrence of an additional transition for the intermediate stiffness regime besides the first-order like aggregation transition ($\gamma > 0$ peak).

into entangled hairpins. In case of 4 polymers, the transitions at lower temperatures are not apparent in the available microcanonical range. Also in the canonical ensemble there are merely small peaks in the corresponding thermal derivatives. This signals that these structural changes at lower energies and corresponding temperatures are second-order like transitions or crossovers. For 8 polymers, there are additional second-order like transition peaks ($\gamma < 0$) visible that may be associated to the transition $A_1 \rightarrow A_2$ from amorphous aggregates into polymer bundles. The structural changes at lower temperatures into twisted bundles are again not visible in the microcanonical ensemble indicating a crossover from polymer bundles to twisted bundles.

It may be noted that the aggregation-peak in $\gamma(E)$ decreases with increasing M which may be overlooked due to the scaling of the axes. This seems surprising remembering that for flexible polymers in Sec. 3.4 one needed larger polymer numbers in order to see a monotonic decrease. This may be due to the lower density and may also change again with intermediate polymer numbers. Still, this supports that also for semiflexible polymer the aggregation transition is a first-order phase transition. To clarify this, Sec. 4.3 includes a microcanonical finite-size scaling analysis.

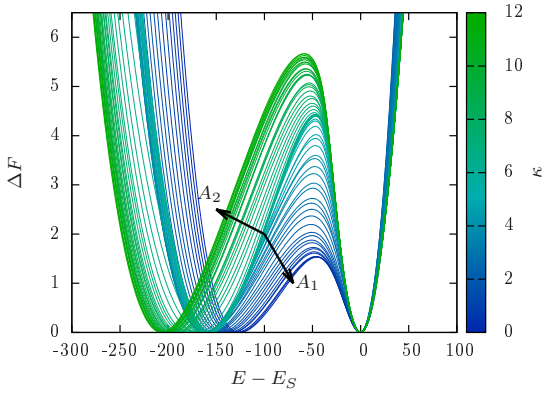

Figure 4.6.: Microcanonical free-energy barrier at the equal-height aggregation temperature depends on the stiffness κ (encoded in the line color as in Fig. 4.3). The energy axes is shifted to the energy of the soluble phase E_S.

4.2.3. Free-energy barrier

Recalling Sec. 3.4, flexible polymer aggregation is a first-order phase transition in the limit of increasing polymer number. Flexible polymers are a limiting case of semiflexible polymers with $\kappa = 0$. Combining this result with the first-oder signatures for finite-size semiflexible polymer aggregation, it is expected that the aggregation transition is accompanied by a free-energy barrier. Figure 4.6 shows an example of the microcanonical free-energy barrier ΔF for 8 polymers, encoding the stiffness in the line color with $\kappa \in [0, 12]$, $\Delta\kappa = 0.2$. The microcanonical free energy is obtained from the canonical energy probability distribution at equal height temperature T_{eqh}:

$$F_{\mathrm{eqh}}(E) = -k_B T_{\mathrm{eqh}} \ln\left(P_{\mathrm{eqh}}(E)\right). \tag{4.5}$$

The free-energy barrier is then defined relative to the free-energy minimum, $\Delta F = F_{\mathrm{eqh}}(E) - F_{\mathrm{min}}$. Alternatively, the microcanonical free-energy may be defined in terms of the transition temperature from the specific heat peak corresponding to an energy probability distribution with both peaks of equal weight. This would emphasize that one phase is more stable than the other, while the overall picture of the barrier is not altered but less clear. In the current definition, the two minima of the free energy correspond to the equilibrium phases at coexistence: the soluble (S) phase at E_S and the aggregated (A) phase at E_A. The existence of a local maximum, or barrier, between these two phases supports the first-order nature already for this finite system size.

Figure 4.6 shows that the free-energy barrier between the coexisting structural phases clearly depends monotonically on the stiffness. In addition, one may qualitatively distinguish the amorphous regime A_1 (blue) from the bundle regime A_2 (green). A finite-size scaling of the barrier with system size follows in Sec. 4.3, but this already supports the recent claim that the free-energy barrier for amorphous aggregation is lower than for aggregation into ordered structures, such as bundles [102].

Figure 4.7.: Parallel symmetric rods (left) and shifted rods (right).

4.2.4. Effective parameterization of Θ-polymer bundles

In order to study the emerging twisted polymer bundles, I make use of effective parameterizations of Θ-polymer bundles at zero temperature. This will not allow to calculate a ground state but to compare different parameterizations and identify possible ground-state candidates. Moreover, the discussion is restricted to bead-spring polymers for which it has been shown that single flexible polymers form icosahedral structures in the ground state [69], similar to Lennard-Jones crystals. In the current parameterization this may be also expected for several flexible polymers. If bending stiffness κ is introduced the situation changes and some possible scenarios are discussed below.

Two parallel rods

The first naive ground state estimate is parallel aligned stiff rods. In this case, the bending potential does not contribute and considering neighboring monomers in their local energy minimum also leads to a vanishing FENE-potential contribution. This leaves only the Lennard-Jones interactions. Assuming for $T \to 0$ a regular structure, Fig. 4.7 shows two examples of parallel stiff rods: symmetric (left) and shifted (right). The shifted case is arranged such that all closest Lennard-Jones contacts are in the minimum r_0, resulting in a $r_0/2$ shift. This can be shown to be the optimal shift. The corresponding total energies for two polymers of length N with a Lennard-Jones cutoff radius $r_c = 2.5\sigma \approx 2.23 r_0 < \sqrt{5} r_0$ are then given by

$$\begin{aligned} E^{\mathrm{p,sym}} &= N V^*_{\mathrm{LJ}}(r_0) + 2(N-1)\left[V^*_{\mathrm{LJ}}(\sqrt{2}r_0) + V^*_{\mathrm{LJ}}(2r_0)\right] \\ &\approx -1.4165N + 0.3350 \\ E^{\mathrm{p,shift}} &= V^*_{\mathrm{LJ}}(r_0)\,(N+N-1) + V^*_{\mathrm{LJ}}(\sqrt{3}r_0)\,(N-1+N-2) + V^*_{\mathrm{LJ}}(2r_0)\,(N-2+N-2) \\ &\approx -2.1095N + 1.2116 \end{aligned}$$

For $N = 13$ this would yield energies $E^{\mathrm{p,sym}} \approx -18.08$ and $E^{\mathrm{p,shift}} \approx -26.21$. This is consistent with the energy scale from simulations.

Two twisted rods

With the estimations of parallel rods in mind, it may be attempted to study the effect of a twist on two semiflexible polymers. Consider two parallel, shifted polymers with monomers located on circles of radius R at distance d, see Fig. 4.8. The polymers are at initial angular distance φ. Now, each polymer gets an additional twist described by an angular change of ω

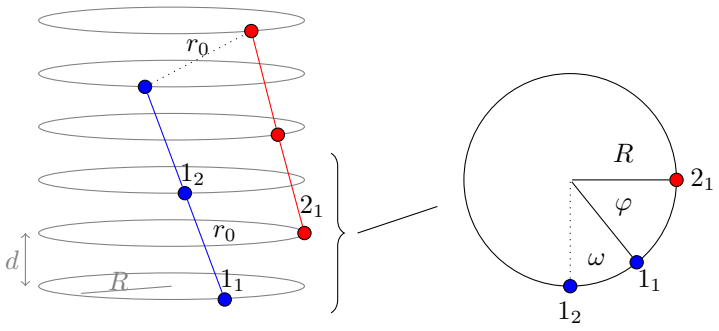

Figure 4.8.: Possible parameterization of 2 twisted rods.

per bond. Again, the distance between neighboring monomers along each polymer is fixed to r_0 as well as the distance between monomers of different polymers with the same index.

This leads to the following constraints on d and R using the law of cosines $c^2 = a^2 + b^2 - 2ab\cos(\gamma)$ in the projected plane of Fig. 4.8 (right):

$$r_0^2 = d^2 + \left[2R^2 - 2R^2\cos(\varphi)\right] \tag{4.6}$$

$$r_0^2 = (2d)^2 + \left[2R^2 - 2R^2\cos(\omega)\right]. \tag{4.7}$$

This is a system of two linear equations with two unknown variables and can therefore be solved yielding

$$R = r_0\sqrt{\frac{3}{2\left(3 - 4\cos(\varphi) + \cos(\omega)\right)}} \tag{4.8}$$

$$d = r_0\sqrt{\frac{\cos(\omega) - \cos(\varphi)}{\cos(\omega) - 4\cos(\varphi) + 3}}. \tag{4.9}$$

For $\omega = 0$ and $\varphi = \pi$, the shifted parallel rods are recaptured with $R = \frac{1}{2}\sqrt{\frac{3}{4}}r_0$ and $d = \frac{1}{2}r_0$. By construction, all angles between neighboring bonds are constant and a function of (d, R).

In principle this may be written in polar coordinates in order to obtain an analytical expression that may be solved numerically. However, it is far more feasible and controllable to use the parameterization in the existing computer program, vary φ and ω and compare the resulting energy including the Lennard-Jones cutoff. Moreover, one only needs to evaluate this one time, measuring the monomer-monomer interaction and the (constant) polar angles separately. This may then be combined with the bending energy in a straight-forward way: Figure 4.9 shows $E_{LJ}(\varphi,\omega) + \kappa 2\ (N-2)\,(1-\cos\theta(\varphi,\omega))$ for selected κ values. It can be seen, that the Lennard-Jones energy of flexible chains may be reduced by the parameterized shift $(1.3\pi, 0.54\pi)$. Of course, this is not compatible with the formation of a spherical globule. If stiffness is introduced, the energetically favorable minimum changes towards $\varphi \to \pi$ and $\omega \to 0$. However, for κ in the regime of polymer bundles there remains an energy minimum with a small twist ($\pi \neq 0$) instead of purely parallel rods.

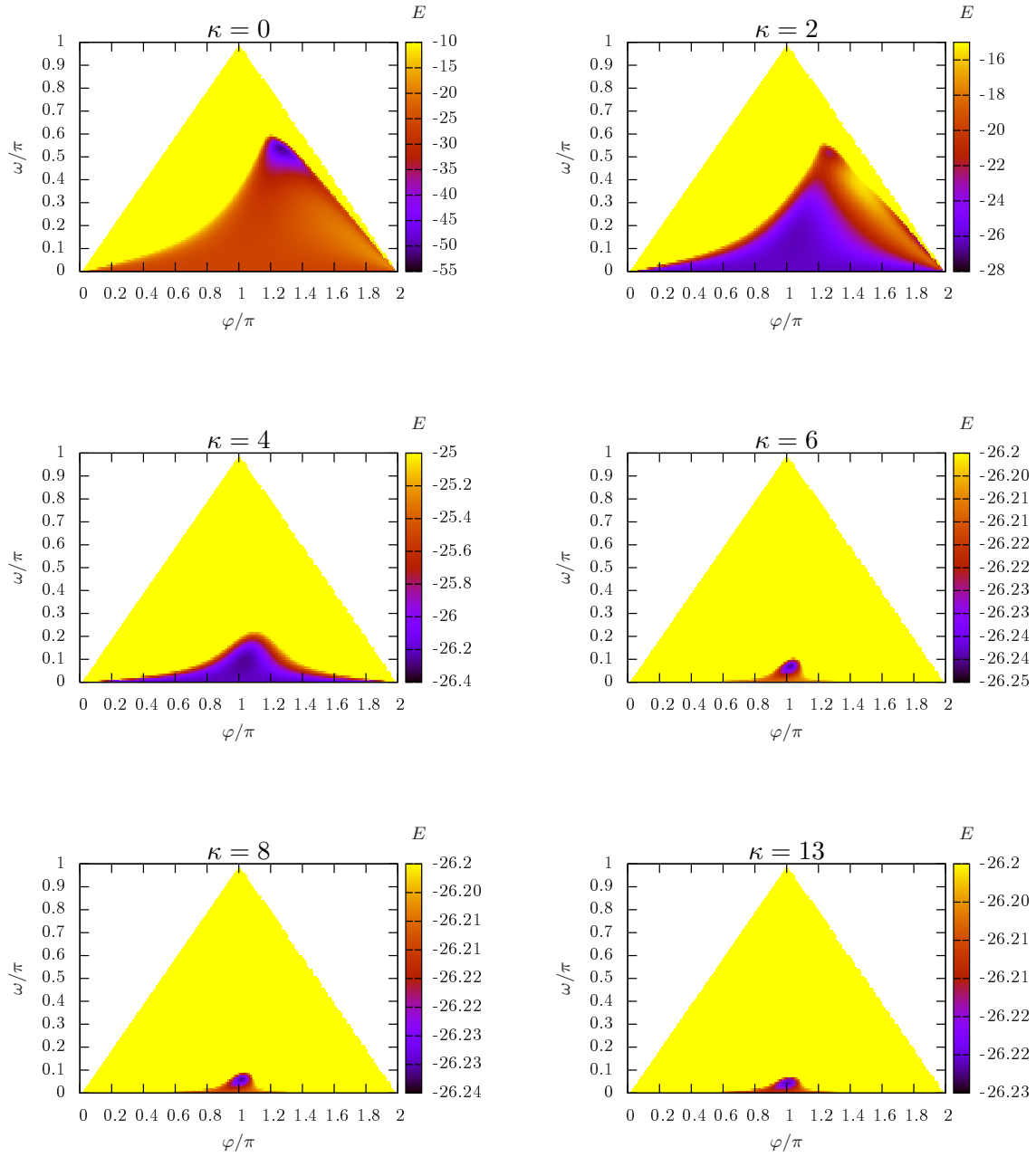

Figure 4.9.: Total energy $E_{LJ}(\varphi,\omega) + \kappa\ 2(N-2)\,(1-\cos\theta(\varphi,\omega))$ from numerical variation of φ and ω for the 2×13 twisted polymers parameterization for several κ.

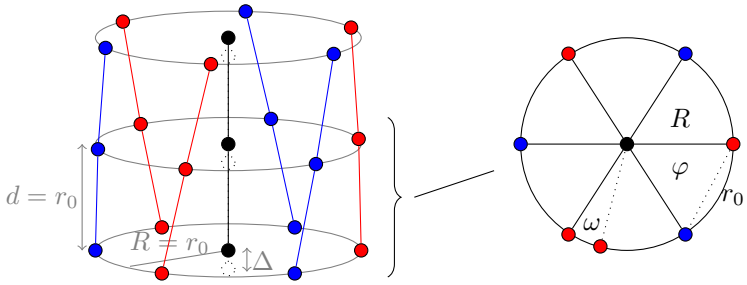

Figure 4.10.: Possible construction of a twisted polymer bundle with 7 polymers.

Twisted bundle - first shell

From the discussion in Sec. 4.2.1 it seems that for more than 2 polymers the low-temperature twisted bundles have a more pronounced twist and form a shell structure. Starting with a bundle core, a hexagonal first shell may be expected as shown in Fig. 4.10. Similar assumptions for polymer bundles have been applied in the literature, e.g., in Refs. [97, 99].

The simplest way to arrange the seven polymers of the first shell is in parallel alignment on a circle of radius $R = r_0$, minimizing all Lennard-Jones contacts due to the hexagonal closed packing with $\varphi = \pi/3$ and $\omega = 0$ in Fig. 4.10. If the monomer layers are at distance $d = r_0$ the FENE contribution vanishes again. Geometric arguments give the following Lennard-Jones contributions

$$\begin{aligned}
&\text{layer } i - \text{layer } i: && 12V^*_{\mathrm{LJ}}(r_0) + 6V^*_{\mathrm{LJ}}(\sqrt{3}r_0) + 3V^*_{\mathrm{LJ}}(2r_0) + \ldots \\
&\text{layer } i - \text{layer } i+1: && 24V^*_{\mathrm{LJ}}(\sqrt{2}r_0) + 12V^*_{\mathrm{LJ}}(2r_0) + 3V^*_{\mathrm{LJ}}(\sqrt{5}r_0) + \ldots \\
&\text{layer } i - \text{layer } i+2: && 7V^*_{\mathrm{LJ}}(2r_0) + 24V^*_{\mathrm{LJ}}(\sqrt{5}r_0) + 6V^*_{\mathrm{LJ}}(\sqrt{7}r_0) + 3V^*_{\mathrm{LJ}}(\sqrt{8}r_0) + \ldots
\end{aligned}$$

Considering a Lennard-Jones interaction with cutoff $r_c = 2.5\sigma \approx 2.23r_0 < \sqrt{5}r_0$, this yields an energy $E^{p,sym,7} \approx -17.700N + 5.6162$ and $E^{p,sym,7}(N = 13) \approx -224.48$. Obviously, this may not be the ground state but further analytical considerations quickly become tedious. Keeping the stiff polymer constraint, Fig. 4.11 (left) shows the total energy from numerical variation of the radius R and a shift Δ of the core polymer, shown by the black, dotted monomers in Fig. 4.10. For a bundle of parallel stiff rods ($\omega = 0$), the optimal conformation is obtained for $\Delta = 0.5r_0$ and $R \approx 0.945r_0$.

While minimizing the full problem of seven twisted rods includes too many free parameters, it is possible to learn qualitative features from a simplified parameterization for the twist. Consider the optimal conformation from the previous parallel bundle. In addition, I introduce a rotation of consecutive monomers along the radius of the given shell by ω. Fixing the outer shell polymer bonds leads to the condition

$$d = r_0\sqrt{2\cos\omega - 1}. \tag{4.10}$$

This will compress the central polymer bond. However, this is affordable since the bond is described by the FENE potential, which shows in the given parameterization less energy increase in the harmonic region than the Lennard-Jones potential. The results are shown in Fig. 4.11 (right). The numerical variation measures the monomer-monomer interaction of

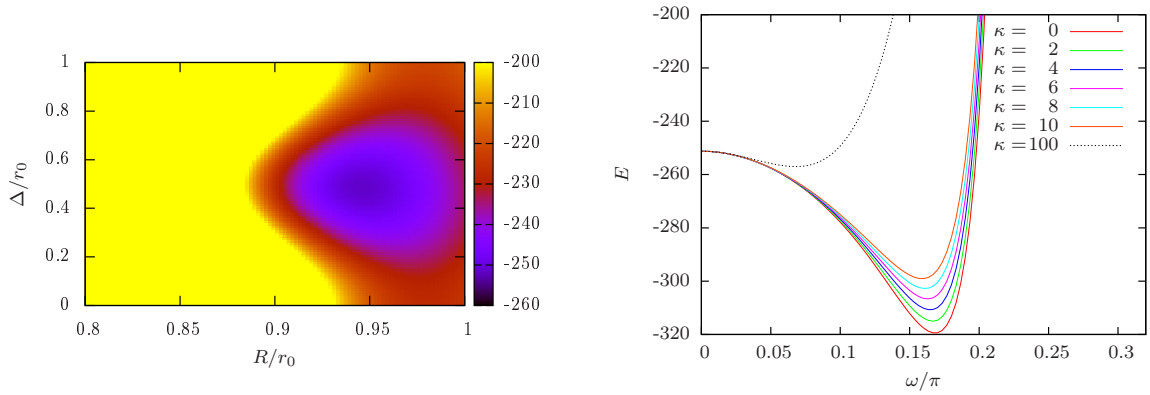

Figure 4.11.: Numerical results for $7 \times N = 13$: (left) Variation of a shift Δ of the core polymer and radius of the first shell R. (right) Variation of the twist ω at the minimal conformation with $\Delta = 0.5r_0$ and $R = 0.945r_0$.

flexible polymers and the (constant) bending angle from the twist. Then, for different bending stiffness one merely needs to linearly add $\kappa \cdot 6(N-2)\,(1-\cos\theta(\varphi,\omega))$ again.

This shows that for the considered parameterization the benefit of the additional Lennard-Jones interactions along different layers outweighs the cost for bending even for large bending stiffness. This may be related to surface-tension arguments, e.g., in Ref [99], and is expected to remain for long and stiff polymer chains. In addition, also flexible polymers show helical ordering in the elongated parameterization, consistent with the results for 2 polymers above and for a single polymer adsorbed onto a nano-wire [101]. This large benefit of (stretched) flexible polymers to form twisted bundles may be exploited: while usually forming amorphous aggregates, one could imagine to artificially stretch flexible polymers, e.g., by an external pulling force or field. The resulting competition between force and twist may allow for a controllable adjustment of bundle twist and corresponding length. This would both be of theoretical and experimental interest, with possible applications in, e.g., molecular measuring devices.

4.3. Finite-size effects depending on polymer stiffness

The above discussion showed first-order like transition signatures for finite semiflexible polymer aggregation. This is not unexpected since the considered model is an extension of flexible Θ-polymers in Chap. 3. Thus, one may expect the same qualitative first-order phase transition behavior in the limit of increasing polymer number $M \to \infty$, as was shown in Sec. 3.4. I consider only semiflexible bead-spring polymers, which limits the accessible system sizes. Therefore, the results presented here will be in the previously noted *intermediate* scaling regime [42]. Also, this reduces the discussion to a merely qualitative comparison with several open questions remaining. For example, the relevant length scale was argued to be the linear extension of the aggregate in Sec. 3.4. This is consistent with particle condensation, because both transition droplet and amorphous aggregate are of spherical shape. On the other hand, as shown above in Sec. 4.2, rather stiff semiflexible polymers form bundles instead. This could influence the geometric arguments made before, depending on the length scale at which the rather stiff polymer aggregates tend to form a spherical (or ellipsoid) structure again. This

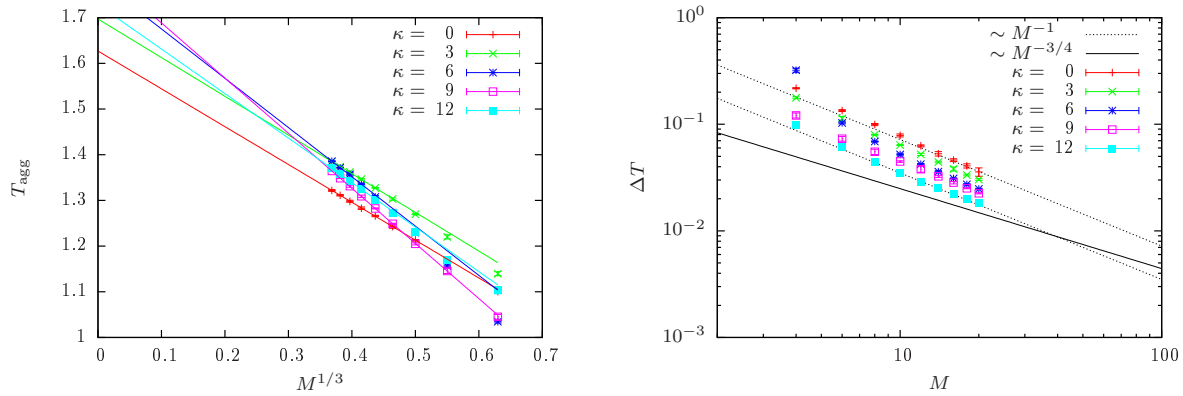

Figure 4.12.: Finite-size dependence of the aggregation temperature (left) and the transition rounding (right) for polymers of length $N = 13$ with selected κ ($\rho_{\mathrm{m}} = 10^{-2}$).

very interesting aspect may not be studied in detail with the available data and has to be left for future investigations. In the following, I will discuss several observables for polymers of length $N = 13$ and monomer density $\rho_{\mathrm{m}} = 10^{-2}$ with selected stiffnesses $\kappa = \{0, 3, 6, 9, 12\}$. I performed parallel multicanonical simulations with up to 256 cores and usually 2.56 million measurements. The maximal number of polymers is limited to 20 in order to have the same data set sizes for all stiffnesses.

The finite-size transition temperature and the finite-size rounding are shown in Fig. 4.12. The aggregation temperature T_{agg} is defined as the temperature where the specific heat has its largest maximum, and the solid lines show the fits with the best (largest) goodness of fit parameter $Q \in [0, 1]$ [82]. All fits include the largest system sizes $M_{\mathrm{max}} \lesssim 20$. As already mentioned this is not very large. Consequently the large-system condensation-like scaling of the form $M^{-1/4}$ may not be expected. However, an intermediate scaling as $M^{-1/3}$ is plausible as observed for flexible polymer aggregation in Sec. 3.4 and Ref. [42]. Since this was geometrically motivated, it may not be trivially generalized to semiflexible polymer aggregation, where bundle formation was observed in Sec. 4.2 and Ref. [95]. Interestingly, the aggregation temperature does not show a deviating behavior with increasing stiffness, which may be partially attributed to the small system sizes. On the contrary, all data yield good fits with the intermediate scaling ansatz. A clear conclusion about a general dependence of the aggregation temperature on the stiffness is not possible.

The finite-size rounding of the transition is obtained from the half-width of the specific heat peak (as in Sec. 2.3.2). Figure 4.12 (right) shows that the rounding ΔT decreases with increasing stiffness. This means that the peaks become sharper with increasing stiffness. The dashed line shows the expected intermediate scaling behavior M^{-1} from particle condensation and flexible polymer aggregation. In addition, the solid line shows the particle-condensation large-system scaling behavior $M^{-3/4}$. The data points seem to follow the expected intermediate scaling regime in all cases. Again, the system sizes are too small in order to yield quantitative conclusions. This is, however, an interesting subject for future investigations. With the present techniques, system sizes of about $M = 60$ should be feasible.

Expecting a scaling behavior similar to particle condensation in the limit of large polymer numbers, another insightful observable is the size of the largest aggregate at the transition.

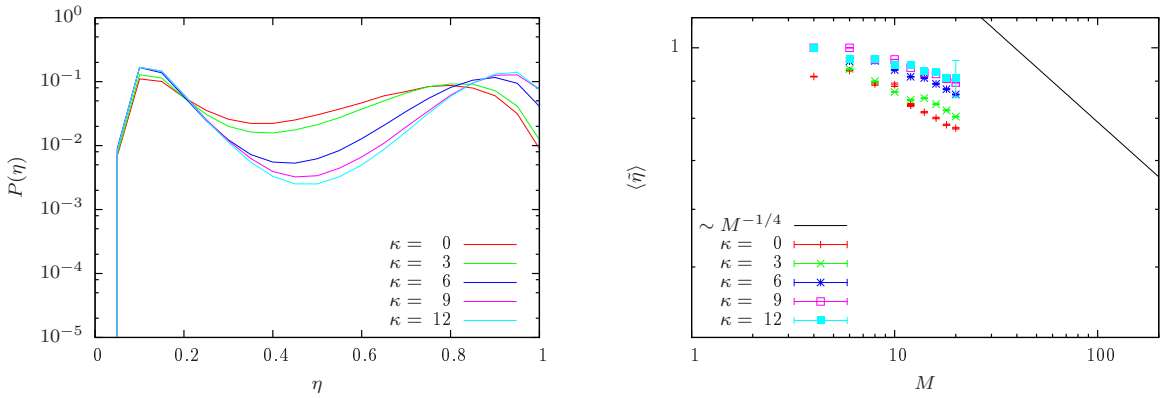

Figure 4.13.: (left) Exemplary aggregate size distributions for $M = 20$ polymers and (right) finite-size dependence of the fraction η of monomers in the largest aggregate at the transition temperature for polymers of length $N = 13$ with selected κ ($\rho_\mathrm{m} = 10^{-2}$).

This may be defined in terms of a monomer fraction $\langle\tilde{\eta}\rangle_{T_\mathrm{agg}}$ in the largest aggregate from Sec. 3.4, considering only the aggregate phase at the transition temperature T_agg. Exemplary distributions of the monomer fraction η are shown in Fig. 4.13 (left) for $M = 20$ polymers and the considered stiffnesses. Relevant for $\langle\tilde{\eta}\rangle_{T_\mathrm{agg}}$ is only the right peak at larger η, i.e., a single aggregate plus unattached polymers. In all cases, the distribution of the largest cluster size is, as expected, similar to the case of flexible polymer aggregation and particle condensation. A peak at large fractions shows that a single macroscopic aggregate forms. With increasing stiffness, however, the right peak tends to even larger fractions. In addition, the suppression of intermediate aggregate sizes is increased. This is also reflected in the average fraction $\langle\tilde{\eta}\rangle_{T_\mathrm{agg}}$, shown in Fig. 4.13 (right). With increasing stiffness, the fraction increases. Importantly, the fraction differs from 1 for larger system sizes, with a decreasing trend. The solid line in the figure shows the expected scaling of the fraction as $M^{-1/4}$, from the balance of surface tension and fluctuation entropy seen for particle condensation. No data set shows this scaling explicitly, which may again be explained by the small system sizes. However, the systematic deviation from 1 may be seen as a first trend towards the expected scaling. More importantly, it supports that a single macroscopic aggregate forms, which may be surrounded by unattached but individual polymers, compare also Sec. 3.4.

The aggregation transition of semiflexible polymers shows similarities to particle condensation on a generic level. This may be identified as the mechanism of phase separation: in both cases a transition occurs between a homogeneous phase and a mixed phase, which consists of a macroscopic cluster plus surrounding constituents in the otherwise homogeneous phase. Then, also semiflexible polymer aggregation should be a common first-order phase transition in the limit of increasing polymer number. Figure 4.14 shows a qualitative microcanonical analysis according to Sec. 2.5. The inflection point analysis [49, 55] allows to classify finite-size phase transitions using the second derivative $\gamma(E)$ of the microcanonical entropy (fixed potential energy). A positive peak in $\gamma(E)$ signals a first-order like transition. If this peak is additionally decreasing towards zero with increasing system size, this is a strong indication of a first-order phase transition in the according thermodynamic limit. This is the case for all

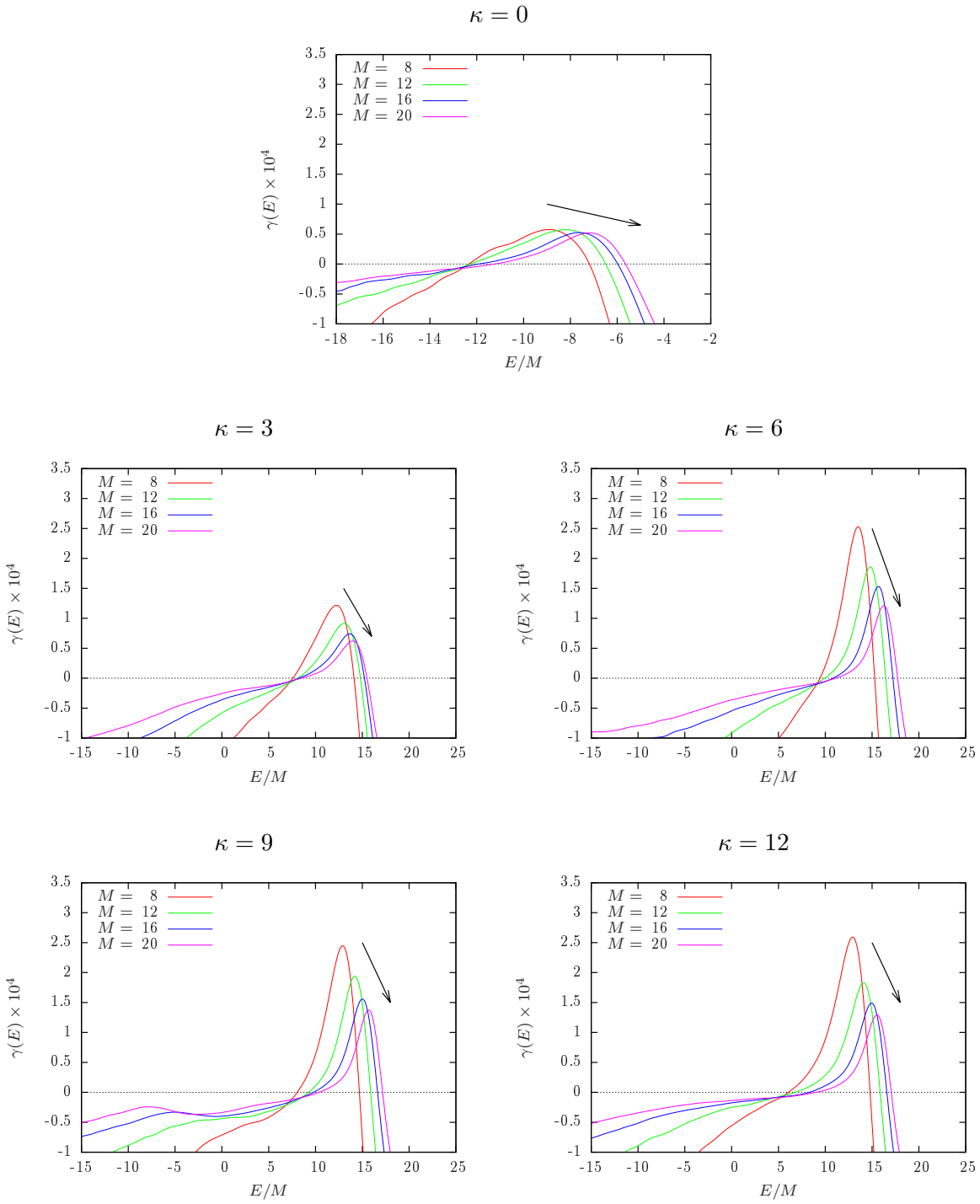

Figure 4.14.: Microcanonical analysis for polymers of length $N = 13$ with selected κ ($\rho_{\mathrm{m}} = 10^{-2}$). A first-order phase transition is characterized by a positive peak in the second derivative $\gamma(E)$ of the microcanonical entropy that approaches zero from above for increasing polymer number (arrows).

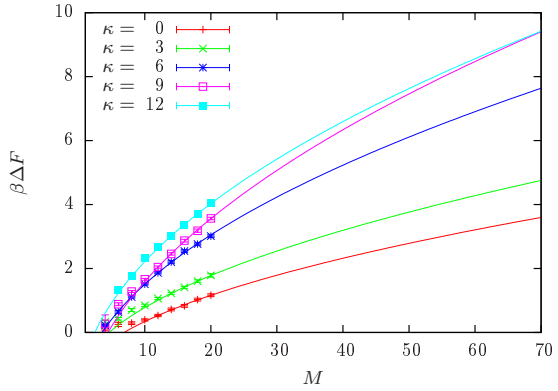

Figure 4.15.: Scaling of the free-energy barrier for selected κ. The barrier increases with increasing stiffness. Solid lines are fits to the condensation-like expected behavior and should be considered rather as guides to the eye due to the small system sizes.

considered stiffnesses. In addition, the peak sizes seem to increase from flexible to rather stiff polymers. For $\kappa = 9$ there are additional second-order like signatures below the aggregation transition (negative peaks in $\gamma(E)$), which are not as systematic but may survive in the considered limit. The presented results show that semiflexible polymer aggregation is a first-order phase transition in the proper limit.

When semiflexible polymer aggregation is a first-order phase transition, it should be accompanied by a free-energy barrier. The free-energy barrier may be estimated from the energy probability distribution $P_{\text{eqh}}(E)$ at the equal-height transition temperature as the ratio between maximum and minimum. Then $\beta\Delta F = (\ln P_{\text{max}} - \ln P_{\text{min}})$, see also Sec. 2.4 and Sec. 3.4. This may be estimated straight-forwardly from multicanonical simulations, obtaining the equal height histograms via binary-search histogram reweighting and estimating the barrier with jackknife errors using histogram reweighting of time-series subsets. Figure 4.15 shows the resulting free-energy barriers for selected κ with increasing number of polymers. It can be seen that the barrier increases with stiffness for all data points, neglecting the fits (solid lines) for now. This is in good agreement with the previously observed increase of the free-energy barrier in Sec. 4.2 and Ref. [95]. Furthermore, it shows that this also holds for increasing system size and should thus remain in the proper finite-size scaling limit. This strongly supports the claim that the free-energy barrier is systematically smaller for amorphous aggregates than for ordered structures, such as bundles or fibrils [102].

Making use of the similarity to particle condensation on a generic level, one may consider a free-energy scaling ansatz according to Eq. (2.63), namely in three dimensions $\beta\Delta F = c_0 + c_1 N^{1/2}$. Of course, there may be additional terms, e.g., a logarithmic contribution considered in Ref. [48]. However, in three-dimensions this simple ansatz described the free-energy barrier scaling for particle condensation already well and is thus considered also for aggregation as solid lines in Fig. 4.15. The (preliminary) fits with optimal goodness-of-fit parameter Q should be considered as guides to the eye, due to the very small system size. However, it may be roughly seen that the slope c_1 increases from amorphous aggregation (small κ, $c_1 \approx 0.6$) to bundle formation (larger κ, $c_1 \approx 1.4$). The numerical data and fits for $\kappa = \{0,\ 3,\ 6,\ 9,\ 12\}$ are

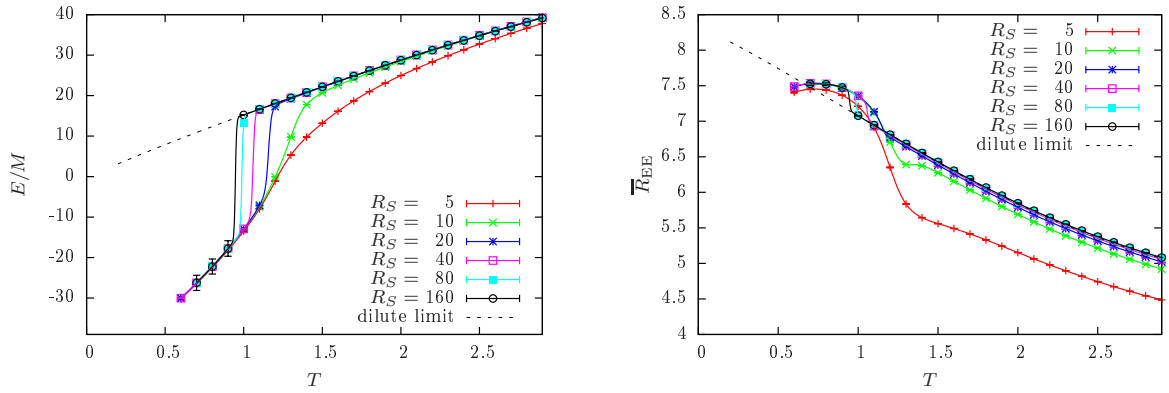

Figure 4.16.: Canonical estimates of the average energy (left) and average end-to-end distance per polymer (right) for $M = 8$ stiff polymers of length $N = 13$ ($\kappa = 9$) in spherical confinement. With increasing radius R_S, the density decreases and the aggregation transition shifts to lower temperatures. The dilute limit refers to a single, isolated polymer.

thus consistent with an increasing free-energy barrier with stiffness as expected. A detailed analysis of this, including larger system sizes, is a promising future investigation.

The conjectured equivalence to particle condensation would require large system sizes, as discussed in Sec. 3.4. This leads to length scales much larger than the polymer extension, i.e., $L \gg Nr_0$. Moreover, the system would be required to be sufficiently dilute. Both conditions are clearly not satisfied by biopolymers (like actin or microtubules) in cells. In fact, the dilute condition is usually not satisfied in real biological systems, which are generally considered to be *crowded*. This emphasises the relevance of the intermediate scaling regime, which should remain dominant when confining biological polymers on typical scales and, in such an approach, may be accessible by experiment.

4.4. Effect of spherical confinement

As an extension to Sec. 3.5, I briefly discuss the effect of density on the aggregation transition of exemplary rather stiff polymers of length $N = 13$ and stiffness $\kappa = 9$. Again, a spherical confinement of radius R_S is considered in order to adjust the density. Parts of this section have been published in Refs. [42, 65].

For the selected parameters and temperature ranges, a single polymer may be considered rather stiff and will not collapse but instead stretch upon a temperature decrease. In this region, the single semiflexible Θ-polymer shows wormlike chain behavior [90]. This single-polymer behavior is shown in Fig. 4.16 as the dilute limit (dashed line) of the average energy per polymer and the average end-to-end distance for 8×13 stiff polymers. For sufficiently large sphere radii, the high temperature behavior is described by the dilute limit. In this regime, the polymers are essentially isolated maximizing the translational entropy within the sphere. The aggregation transition results in a sharp drop of the average energy and an initial additional stretch of the individual polymers. Below the aggregation transition, the low-temperature behavior is again independent on the density: with decreasing temperature,

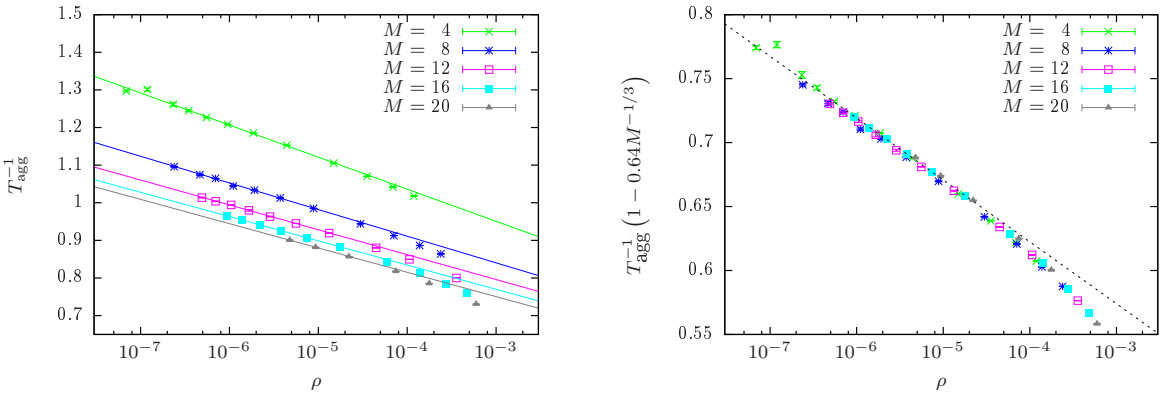

Figure 4.17.: Density dependence of the inverse aggregation temperature for stiff ($\kappa = 9$) polymers of length $N = 13$ in spherical confinement with fits to sufficiently low densities (left). Rescaling the inverse temperature with finite-size scaling expectations from Sec. 4.3 allows for a data collapse.

the average energy monotonically decreases while the average end-to-end distance first increases and then decreases as well. In both cases, the energy is minimized by maximizing polymer contacts - first by stretching in order to align within polymer bundles, and afterwards by forming twisted bundles at lower temperatures as observed before in Sec. 4.2. This density depending isolated-chain versus bundle-formation behavior is consistent with experiments and simulations of (rather stiff) DNA solutions [107]

As observed for flexible polymers in Sec. 3.5.1, the high-temperature dilute regime of individual, separated polymers is only achieved for sufficiently dilute systems. For flexible polymers, the onset of the semi-dilute regime was estimated comparing volume-fraction estimates by de Gennes [3] of a multi-chain and a single-chain system, see Eq. (3.15). While this may not be completely generalized to stiff polymers, it seems to be able to predict the order of magnitude as well. For 8 flexible polymers of length $N = 13$, the onset yields $R_S^c \approx 6.3$ assuming a self-avoiding walk exponent. Considering for stiff polymers instead a maximal exponent of $\nu = 1$, leads to an upper bound of $R_S^c \approx 18.2$, which is in decent agreement with the onset of the dilute regime, as observed in Fig. 4.16.

The density dependence of the aggregation temperature, estimated in Sec. 3.5.2 from micro-canonical arguments comparing the conformational entropy of the aggregated and the "ideal-gas-like" separated phase should still remain valid for stiff polymers. This means that the inverse aggregation temperature should be linearly depending on the logarithm of the polymer density, i.e., $T_{\mathrm{agg}}^{-1}(\rho) = -c_1 \ln \rho + c_2$, see also Sec. 2.6. This is supported by results for specific heterogeneous lattice polymers as a model for protein aggregation [60]. In addition, Fig. 4.17 (left) shows exemplary results for the inverse aggregation temperature for up to $M = 20$ rather stiff bead-spring homopolymers of length $N = 13$ at different densities. This clearly confirms the expected density dependence for sufficiently low densities. Notice that "sufficiently small" seems to be over an order-of-magnitude smaller for stiffer polymers than for flexible polymers, where for the same polymer length densities $\rho < 10^{-3}$ were described by the logarithmic dependence.

As for the flexible polymers, I apply the finite-size scaling arguments from Sec. 4.3 for semiflexible and rather stiff polymers in order to rescale the inverse temperature axis with the extrapolation of the intermediate scaling regime. To this end, the density dependence of the leading-order correction amplitude is neglected and an effective value chosen instead. The resulting data collapse is shown in Fig. 4.17 (right) with the rescaling parameters given in the axis label. This emphasizes the onset of sufficiently small densities for the logarithmic density dependence. Furthermore, it shows that the finite-size effects in the intermediate regime allow for a stable description of finite-size dependence. Compared to Sec. 2.6 and Sec 3.5.1, this highlights the generic density dependence of these phase separation transitions, caused by a reduction of translational entropy from homogeneous to mixed phases.

5. Computational methods

The previous numerical results mostly rely on Markov Chain Monte Carlo simulation methods. While loosing any information about the dynamics, these methods allow to sample the configuration space with a given ensemble probability. The most-known example is the Metropolis algorithm [111] in the canonical ensemble. If several canonical ensembles are simulated in parallel with exchange of configurations, we talk about the parallel tempering method [112–116]. However, also other ensembles may be considered, for example the microcanonical ensemble [49, 50, 56–59] or the "multicanonical" ensemble [117–119]. For the discrete systems, I additionally applied exact enumerations in order to validate the numerical results. While most of the results are produced with a parallel version of the multicanonical method [120], I usually applied a combination of algorithms. This will be discussed in the following sections, where I will shortly describe the underlying concept and the applied methods with a focus on the parallel multicanonical simulation.

5.1. Monte Carlo simulations in the canonical ensemble

In order to understand the concept of enumeration, importance sampling and reweighting, it is unavoidable to consider the formulation of statistical mechanics in the canonical ensemble (for detailed discussions see, e.g., Refs. [6, 121–123]). Here, one considers a system in equilibrium with a thermal reservoir at a given temperature T or inverse temperature $\beta = (k_\mathrm{B}T)^{-1}$, i.e., there is no net macroscopic flow of matter or energy. Consider the phase space spanned by the momentum vector p and the position vector q, describing a complete classical system. The Hamiltonian $\mathcal{H}(p,q)$ encodes all interactions in the system, relating momentum and position to an energy. The canonical probability density in phase space is given by

$$P(p,q) = \frac{1}{Z(T,V,N)}\, e^{-\beta\mathcal{H}(p,q)}, \tag{5.1}$$

normalized with the partition function over all possible states

$$Z(T,V,N) = \int dp \int dq\; e^{-\beta\mathcal{H}(p,q)}. \tag{5.2}$$

Now, for any system in which the forces do not depend on the momenta we can explicitly integrate over the momenta. While this is not possible for systems with moving particles in a magnetic field, it is applicable for most systems of interest. Then, for classical systems, the Hamiltonian can be merely written as a contribution from the kinetic energy and the potential

energy and for N constituents in d dimensions we end up with

$$\begin{aligned} Z &= \int dp \int dq \ \exp\left[-\beta\left(\sum_{i=1}^{dN}\frac{p_i^2}{2m} + V(q)\right)\right] \\ &= \int dp \ \exp\left[-\beta\sum_{i=1}^{dN}\frac{p_i^2}{2m}\right]\int dq \ e^{-\beta V(q)} \\ &= (2\pi m k_\mathrm{B}T)^{dN/2}\int dq \ e^{-\beta V(q)}. \end{aligned} \tag{5.3}$$

This allows to only consider the configuration space, which is in fact what one usually does in classical Monte Carlo simulations. Ignoring the time trajectories of a system, Monte Carlo simulations in classical statistical physics sample the canonical equilibrium probability distribution

$$P(q) = \frac{1}{Z_\mathrm{MC}} \ e^{-\beta V(q)}, \tag{5.4}$$

with the normalization

$$Z_\mathrm{MC} = \int dq \ e^{-\beta V(q)}, \tag{5.5}$$

which may be considered as a Monte Carlo (MC) partition function. This may be rewritten with respect to any suitable variable (or in fact several variables) of choice, for example the potential energy E

$$Z_\mathrm{MC} = \int dE \ \Omega(E) \ e^{-\beta E}, \tag{5.6}$$

where $\Omega(E)$ is called the density of states.

In practice, the integrations usually will be sums either due to discrete systems, artificial discretization of continuous systems or latest due to discretization in the computer program. Moreover, it is not possible to sample every state of the system in a finite amount of time. Instead, Markov Chain Monte Carlo simulations make use of *importance sampling*, such that in stationary equilibrium "detailed balance" is fulfilled. If $p(q \to q')$ denotes the transition probability from state q to state q', this means that the master equation is time independent, i.e. $\frac{d}{dt}P(q,t) = 0 = \int [P(q')p(q' \to q) - P(q)p(q \to q')]dq'$. This is solved trivially by the "detailed balance" condition

$$P(q)p(q \to q') = P(q')p(q' \to q), \tag{5.7}$$

as one possible solution. Now, the transition probability is the product of the probability to select a new state $p^s(q \to q')$ and to accept it $p^a(q \to q')$. Then, one solution to Eq. (5.7) is the following rule for the acceptance probability

$$p^a(q \to q') = \min\left(1, \frac{P(q')p^s(q' \to q)}{P(q)p^s(q \to q')}\right). \tag{5.8}$$

For symmetric selection probabilities, $p^s(q \to q') = p^s(q' \to q)$, this yields the famous *Metropolis algorithm* [111]: With the Boltzmann weight $P(q) \propto \exp\{-\beta E(q)\}$, the expression simplifies to $\min(1, \exp\{-\beta\Delta E\})$. This means that a trial configuration is generated and immediately accepted if $\Delta E \leq 0$ or else accepted with probability $\exp\{-\beta\Delta E\}$.

The resulting importance sampling leads to estimators of the expectation value in configuration space. I want to briefly discuss the difference between the canonical expectation value of an observable $\langle O \rangle$ and the corresponding estimator $\overline{O}$. The expectation value is the integral (or the sum) over the full configuration space $\{q\}$, weighting the observable with its probability weight $P(q)$ and is exact:

$$\langle O \rangle = \frac{1}{Z_{\mathrm{MC}}} \int dq \; O(q) \; e^{-\beta E(q)}. \tag{5.9}$$

Exact enumerations (see next section) yield the expectation value. However, using equilibrated Markov chains allows to estimate the expectation value as its arithmetic mean [6]

$$\langle O \rangle = \int dq \; O(q) P(q) \approx \overline{O} = \frac{1}{N} \sum_{k=1}^{N} O(x_k). \tag{5.10}$$

As mentioned above, one often only considers the sum due to discretization. Then, the above equations change according to

$$\int dq \to \sum_{\{q\}} \; . \tag{5.11}$$

In general, I will obtain most relevant information from the thermal derivatives of observables. Strong changes in observables, associated with phase transitions, are apparent as a peak in the first derivative and a zero crossing in the second derivative. One prominent example is the specific heat $\langle C_V \rangle = \frac{d}{dT} \langle E \rangle / V$. Considering the definition of the expectation value in the canonical ensemble Eq. (5.9), the thermal derivative may be computed

$$\frac{d}{dT} \langle O \rangle = \beta^2 \left(\langle OE \rangle - \langle O \rangle \langle E \rangle \right). \tag{5.12}$$

This may be generalized in terms of estimators, yielding

$$\frac{d}{dT} \overline{O} = \beta^2 \left(\overline{OE} - \overline{O} \; \overline{E} \right). \tag{5.13}$$

5.2. Exact enumerations

The exact enumerations performed throughout this thesis are simple recursive sums over the full state space $\{q\}$ of discrete systems. The idea is to perform the sum of Eq. (5.6) explicitly and hence enumerate the density of states $\Omega(E)$, or possibly $\Omega(X)$ where X is a set of variables. In order to evaluate observables which are not constant for fixed X, one needs to additionally consider the sum of each observable over the full space with fixed X: $O_\Sigma(X) = \sum_q O(q) \delta_{X(q)X}$. Note that $\Omega(X) = 1_\Sigma(X)$ is in this sense a "counting observable". This allows to calculate exact canonical expectation values at any temperature with

$$\begin{aligned} Z_{\mathrm{EE}} &= \sum_X \Omega(X) \; e^{-\beta E(X)} \\ \langle O \rangle_\beta &= \frac{1}{Z_{\mathrm{EE}}} \sum_X O_\Sigma(X) \; e^{-\beta E(X)}. \end{aligned} \tag{5.14}$$

In case of the Ising model, this would mean to consider a very small, initially homogeneously filled, lattice with N spins and recursively flip single spins until all 2^N conformations are generated. Upon a spin flip, the magnetization ($\sum s_i$) changes. This corresponds to the canonical ensemble of the Ising model. However, in this thesis I considered in a first step gas systems with M particles equivalent to the Ising model at fixed magnetization, see Sec 2.1.1. Then, the enumeration has to generate a *fixed* number of particles on a small lattice. This can be done recursively as follows: start by putting a single particle successively on each site of the lattice. Now, for every realization of a particle on a site, one repeats this procedure with a second particle on all remaining available sites. Again, for every combination of two particles on the lattice, one successively puts a third particle on each available site and so on. Every time when the last particle is added the energy as well as other possible observables get evaluated and counted in the density of states and corresponding arrays, respectively. Thus, I was quickly able to enumerate up to 8 particles on a two-dimensional 7×7 square lattice with periodic boundary conditions, which suffices for comparing purposes.

In the case of lattice polymers, this was a little more demanding. The single interacting self-avoiding polymer of length N (number of monomers) may be also enumerated with a recursive function. Consider a polymer conformation already at length $N-1$, then all non-occupied neighbors to the $(N-1)^{\text{th}}$ monomer get successively occupied one after another, always evaluating the observables of interest such as the energy, the number of contacts, the end-to-end distance and so on. The polymer conformation of length $N-1$ itself is only one of the generated ones from a previous conformation of length $N-2$ and so on. That way, all possible combinations of polymer conformations of length N are generated and evaluated. In three dimensions, I was able to enumerate a single polymer of length $N = 18$ without additional tricks on a conventional CPU within a day. Of course, applying symmetry considerations would allow longer chains or shorter computing times, but would not allow to add another polymer easily.

As I am ultimately interested in polymer aggregation, I had to combine this with the previous method for particle condensation. The first polymer is enumerated with a fixed starting point. For each complete configuration (N-monomers), a second polymer is enumerated at a second fixed starting point. In principle, this procedure has to be repeated for all pairs of sites on the lattice. However, considering only two polymers on a cubic lattice of size $L > 2N+1$ (periodic boundary conditions), one can make use of the following symmetry considerations: the first polymer may be fixed at the center due to translational invariance. The second starting point may be chosen from a 16th of the volume, namely half of any of the octants, due to rotational and mirror symmetries. Every starting point then gets a factor of 16, unless on the boundaries for which less mirror points exist. Moreover, I distinguished between inter-chain contacts n_i and intra-chain contacts n_o, sorting the sum of observables in two-dimensional arrays $O_\Sigma[n_i][n_o]$. This allows to reweight to the canonical expectation value depending on temperature and interaction ratio $\epsilon' = \epsilon_o/\epsilon_i$ with $E(X) = \epsilon_i n_i + \epsilon_o n_o$

$$
\begin{aligned}
Z &= \sum_{n_i,n_o} \Omega(n_i,n_o)\; e^{-\beta(\epsilon_i n_i - \epsilon_o n_o)} \\
\langle O \rangle_{\beta,\epsilon'} &= \frac{1}{Z} \sum_{n_i,n_o} O_\Sigma[n_i,n_o]\; e^{-\beta(\epsilon_i n_i - \epsilon_o n_o)}
\end{aligned}
\tag{5.15}
$$

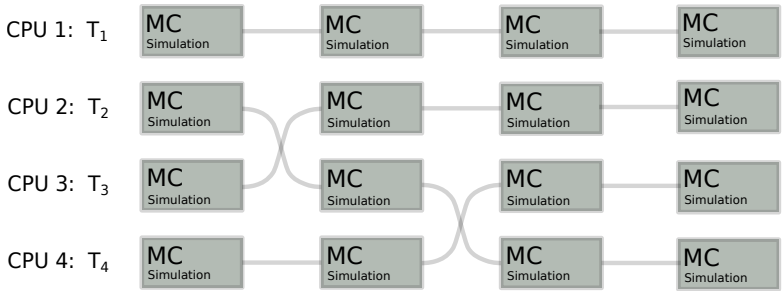

Figure 5.1.: Scheme of parallel tempering taken from Ref. [124]. The system is simulated for a set of temperatures in parallel with the possibility that two Markov chains may exchange configurations according to Eq. (5.17).

5.3. Parallel tempering

Parallel tempering [112–116] uses a number of replica at different temperatures that may interchange system configurations. It is also known as replica-exchange Markov Chain Monte Carlo sampling and is based on the Metropolis algorithm and the general considerations in Sec. 5.1. In principle, p copies of the systems are simulated simultaneously at different inverse temperatures $\beta_1 < \beta_2 ... < \beta_p$. The temperatures may be, in the easiest way, equally distributed in a desired interval $T_i \in [T_{\min}, T_{\max}]$ or more complex schemes may be applied. The important requirement is that the resulting histograms $H(E)$ overlap. These copies are randomly initialized and simulated with the Metropolis algorithm [111] individually, but at fixed update intervals the copies may exchange configurations with a certain acceptance probability maintaining equilibrium, see Fig. 5.1. The idea is to circumvent the slow Monte Carlo dynamics at low temperature: if a system gets stuck in a (low-temperature) local energy minimum it may now change back into a higher-temperature replica and "walk around" possible barriers. As mentioned before, the replica exchange should maintain equilibrium and hence each fulfill ergodicity and detailed balance according to Eq. (5.8). Consider two states q and q' at inverse temperature β and β' respectively. Since the two replica may be considered independent, the joint probability is the product of the individual probabilities

$$P(q, q') = \frac{1}{ZZ'}\ e^{-\beta E(q)}\ e^{-\beta' E(q')}. \tag{5.16}$$

Choosing only updates with symmetric selection probabilities, i.e. $p^s(qq' \to q'q) = p^s(q'q \to qq')$, the acceptance probability simplifies to

$$p^a(qq' \to q'q) = \min\left(1, \frac{e^{-\beta E(q) - \beta' E(q')}}{e^{-\beta' E(q) - \beta E(q')}}\right) = \min\left(1, e^{-\Delta\beta \Delta E}\right) \tag{5.17}$$

That way, emerging barriers in the free-energy landscape may be "walked around" in principle. However, this requires an appropriate overlap of neighboring energy probability distributions which may become quite cumbersome. In this thesis, I am merely using the parallel tempering method for small system sizes in order to estimate the energy range for multicanonical simulations. Thus, the temperatures are equally distributed from $T_{\min}$ to $T_{\max}$ and a suitable number of sweeps and replica changes are chosen such that the system relaxes to the desired

temperature range. In the end, I consider the average energies in the lowest-temperature and the highest-temperature replica for further processing.

5.4. The multicanonical method

The multicanonical method (MUCA) [117–119] allows to sample a broad parameter space with a single simulation. This is achieved by modifying the acceptance probability of updates such that the simulated distribution becomes flat over the full sampling range, hence often also referred to as "flat-histogram method". Previously suppressed states have to become more probable. To this end, the Boltzmann factor in the probability distribution $P(q) \propto \exp\{-\beta E(q)\}$ is replaced, or sometimes extended, by a weight function $W(E(q))$. The partition function may thus be rewritten as

$$Z_{\text{can}} = \int dq\ e^{-\beta E(q)} = \int dE\ \Omega(E) e^{-\beta E} \tag{5.18}$$

$$\rightarrow Z_{\text{MUCA}} = \int dq\ W\left(E\left(q\right)\right) = \int dE\ \Omega(E) W(E). \tag{5.19}$$

This is formulated here in terms of the (potential) energy E, because it is the suitable choice, if one wants to reweight to a set of canonical ensembles or temperatures in the end. Moreover, this may be applied in generalized formulations leading to multimagnetic, multibondic or even other realizations. For this thesis, my desired ensemble is the canonical ensemble and I stick to the (more intuitive) formulation in terms of the energy E. The acceptance probability for symmetric selection properties is then given by

$$p^a(E \rightarrow E') = \min\left(1, \frac{W(E')}{W(E)}\right) \tag{5.20}$$

The most demanding part is to iteratively modify $W(E)$ in order to yield a flat histogram. From Eq. (5.19) it can be seen that a flat histogram - meaning all energies from a chosen interval have the same probability or contribution to the partition function - requires $W(E) \approx \Omega^{-1}(E)$. So in principle, one needs to estimate the density of states. That is a common issue in statistical physics, but also shows the potential of the method or equivalent methods like Wang-Landau [125, 126]. The general procedure is to modify $W(E)$ until a flat histogram is obtained and, in the end, to perform a *production run* with fixed weights sampling the desired statistics. While the Wang-Landau method modifies the weight function after every update, the general procedure in multicanonical simulations is to perform equilibrium simulations at fixed weights in each iteration and to modify the weights in between iterations. In Sec. 5.4.2, I will discuss two possibilities to update the multicanonical weight function and additional modifications that I applied. As mentioned above, the multicanonical weights are related to the density of states which usually covers several hundred or even several thousand orders of magnitude. It is thus in any case advantageous to consider logarithmic weights $\ln(W(E))$ for numerical reasons.

Formally, the above formulation is valid for a continuous energy space. However, due to numerical reasons, I consider an equally discretized energy space on which I define the histograms and weight function. For the lattice systems I considered, this is the exact energy space, while for off-lattice systems energies are sorted into bins $\left[E_i - \frac{\Delta E}{2}, E_i + \frac{\Delta E}{2}\right)$. For the artificially

discretized systems a definition of canonical expectation values via histograms on this discrete space as in Eq. (5.14) would yield systematic deviations, since the conformations are exponentially suppressed within each bin. However, I only apply the discretized space to the weight function and the histograms but not to the measured observables in the final production run. This allows to apply standard reweighting techniques to obtain precise results.

5.4.1. Reweighting to other ensembles

In general, having sampled data according to a multicanonical distribution $W(q)$, i.e., in a (fixed) multicanonical ensemble, allows to reweight to a desired distribution $P(q)$ in other ensembles using standard reweighting techniques. As usual, this requires that the desired distribution is covered by the available data and that throughout the *production run* the multicanonical weights are fixed. Now, the estimator of an expectation value in the desired ensemble may be reweighted from importance sampled multicanonical data by performing the ensemble average, where each contribution is multiplied with the desired probability $P(q)$ and divided by the sampled probability $W(q)$:

$$\langle O \rangle = \frac{\left\langle OP(q)W^{-1}(q) \right\rangle_{\mathrm{MUCA}}}{\left\langle P(q)W^{-1}(q) \right\rangle_{\mathrm{MUCA}}}. \tag{5.21}$$

Here, $\langle \dots \rangle_{\mathrm{MUCA}}$ refers to the ensemble average in the multicanonical ensemble.

Reweighting to the *canonical ensemble*, the desired distribution is the Boltzmann distribution $P(q) = P(E(q)) = \exp\{-\beta E(q)\}$. This allows to express the above equation in terms of the energy and serves as a good example for the following reweighting techniques applied throughout this thesis. Since most systems considered in this thesis include continuous variables (especially the energy), this requires *time-series reweighting* for non-biased estimates of expectation values. As the name suggests, the time series (measurements at fixed intervals) from the production run is directly used and with it the full precision of the measured observables. The concept is the same as above: the multicanonical ensemble average is modified such that the sampled probability is "removed" and the desired probability is "added":

$$\langle O \rangle_{\mathrm{can}} = \frac{\frac{1}{N} \sum O_i \; e^{-\beta E_i}/W(E_i)}{\frac{1}{N} \sum e^{-\beta E_i}/W(E_i)}. \tag{5.22}$$

This approach is as systematic as possible, since each precisely measured value is weighted with the precise probability, divided by the possibly discrete weight function. The discrete weight function is not a problem as long as it is fixed throughout the production run (importance sampled according to $W(E)$), because it is subtracted out in the end.

Another technique is *histogram reweighting*, which may introduce a (small) systematic bias for continuous variables, depending on the bin size of the discrete histograms. In general, the time series of the production run would be sorted into a histogram. For discrete systems it is often often possible to record the histogram directly within the simulation. In these cases (e.g. lattice gas and lattice polymers) and according bin size, this remains as precise as time-series reweighting. In any case, it allows to quickly reweight to the canonical probability distribution

$$P_{\mathrm{can}}(E) = \frac{H(E)e^{-\beta E_i}/W(E_i)}{\sum_E H(E)e^{-\beta E_i}/W(E_i)}, \tag{5.23}$$

and also to estimate canonical expectation values. In the latter case, the observable is usually not constant throughout an energy interval such that the sum of observables at equal energy $O_\Sigma(E) = \sum_{E'} O(E')\delta_{EE'}$ is considered, as in the case of exact enumeration (see Eq. (5.14)). Again, for continuous energies this may be obtained from the times series by sorting into a binned histogram. The data from the production run is then reweighted according to

$$\langle O \rangle_{\text{can}} = \frac{\sum_E O_\Sigma(E) e^{-\beta E_i}/W(E_i)}{\sum_E H(E) e^{-\beta E_i}/W(E_i)}. \tag{5.24}$$

This is a lot faster than time-series reweighting at the cost of a systematic bias when weighting variables with continuous energies in the interval $\left[E_i - \frac{\Delta E}{2}, E_i + \frac{\Delta E}{2}\right)$ with the same weight.

Another important ensemble considered in this thesis is the *microcanonical ensemble*, which may provide complementary information when studying transitions [49, 50]. Moreover, a multicanonical simulation in terms of the (potential) energy E is already somehow in the microcanonical ensemble as all functions depend on E. Here, the microcanonical ensemble is defined as the set of states with a fixed (potential) energy E. As mentioned before, the weight function is approximately proportional to the inverse density of states related to the microcanonical entropy: $W(E)^{-1} \approx \Omega(E) = \exp\{S(E)\}$. It is possible to estimate the microcanonical entropy from the histogram of the final production run, such that $S(E) = \ln(H(E)) - \ln(W(E))$. From this, we may obtain the microcanonical caloric temperature $\beta(E) = \frac{\partial S}{\partial E}$, as well as the second derivative $\gamma(E) = \frac{\partial^2 S}{\partial E^2}$, which is just the inverse microcanonical specific heat $\gamma(E) = -\left[C_{\text{micro}}(E)/\beta(E)^2\right]^{-1}$ [95]. For qualitative statements this and the numerical derivatives

$$f'(E) \approx \frac{-f(E+2\Delta E) + 8f(E+\Delta E) - 8f(E-\Delta E) + f(E-2\Delta E)}{12\Delta E}, \tag{5.25}$$

$$f''(E) \approx \frac{-f(E+2\Delta E) + 16f(E+\Delta E) - 30f(E) + 16f(E-\Delta E) - f(E-2\Delta E)}{12\Delta E^2}, \tag{5.26}$$

are sufficient, while having truncation errors and introducing the previously mentioned small bias. This still allows for a qualitative microcanonical analysis [49–51, 55]. In case of data with large noise one may apply Bézier curves [127]. For an example see Sec. 2.5. When a precise estimate and a proper error estimation is required, then the reweighting should be performed analogue to the time-series reweighting in the canonical case described above.

It needs to be mentioned that for the microcanonical ensemble one may distinguish between two cases. A practical choice from Monte Carlo simulations is to consider constant potential energy E, such that the weight becomes the Kronecker-Delta function $\delta_{E'E}$. The formal definition, however, is a fixed total energy $E_t = E + E_k$ as the sum of potential energy E and kinetic energy E_k. Remembering that we integrated over all momenta Eq. (5.3), this fixed-energy constraint introduces a Θ-function and results in a weight depending on the degrees of freedom (dof), namely $(E_t - E)^{(\text{dof}-2)/2}$, which replaces the Boltzmann weight [56–59]. For a Monte Carlo simulation in configuration space, the kinetic energy is then merely a "reservoir" where additional energy may be stored in or taken from.

5.4.2. Weight modification

The simplest iteration scheme is to estimate the consecutive weights from the current histogram via the *trivial modification* $W^{(n+1)}(E) = W^{(n)}/H(E)$. This corresponds to a local estimation

of the density of states $\Omega(E) \approx H(E)/W(E)$, which is used in the estimation of the consecutive weight function. The weight function, on the other hand, is assumed to sample the inverse density of states $W^{(n+1)}(E) \approx \Omega^{-1}(E)$, leading in a self-consistent approach to the update rule. One immediately sees that states with a high probability (large histogram entries) get a lower probability in the consecutive weights, while suppresses states get a relatively higher probability. In general, this is a good starting point and might even outperform "smart" modifications. The problem for this trivial modification is, however, that it needs a large amount of samples in each single iteration in order to work reliably.

The most intuitive and simple approach is to begin with weights according to the canonical ensemble at infinite temperature, i.e. $\beta = 0$ and $W = const.$ However, it may be suitable to start at a finite temperature T_0, where $W(E) = \exp\{-E/T_0\}$, to save computational time in the iterations. Especially for first-order like transitions it is beneficial to systematically adjust the weights from high to low energies, or vice versa. If starting at high energies, then the initial temperature should be sufficiently above the expected transition temperature.

Another possibility to speedup the convergence is to extrapolate the logarithmic weights at the boundaries of the so far sampled region. For this, it is crucial to have a rough understanding of the weight functions final form. Consider for example the lower energy bound. After the i^{th} iteration, the logarithmic weights have been modified down to E^i_{low}. It may be expected that the logarithmic weights will be monotonically increasing with decreasing energy (this however strongly depends on the studied system and also on E^i_{low}). A rather safe choice is to extrapolate the logarithmic weights with the Boltzmann weight of a sufficiently large temperature, say T_{max}. Moreover, one might measure the local slope of the modified weight function above E^i_{low} and extrapolate this into the range below E^i_{low}. At first-order phase transitions this needs additional care due to the "backbending effect" [49] in the microcanonical temperature and the corresponding so-called "convex intruder" in the microcanonical entropy [51–54]. This convex intruder in the logarithmic weight function (the negative microcanonical entropy) may cause a local slope smaller than the convex hull (see also Sec. 2.5) which may over-estimate the required weight for low-energy states. In the worst case, this may lead to an amplification of the low-energy regime which may become lengthy to undo by the considered update algorithm. The same may be applied to the upper bound analogously.

As mentioned above, the "trivial modification" requires a large amount of samples per sweep that is forgotten afterwards. A possible way out is the *recursive weight modification* [119]. This scheme considers all previous iterations in the weight update using a statistically weighted average. From the general definition Eq. (5.19) it is evident that each iteration locally allows to estimate the density of states $\Omega(E) \approx H(E)/W(E)$. Since the microcanonical entropy $S(E)$ is defined as the logarithm of the density of states $S(E) = \ln(\Omega(E))$, it is convenient to express the consecutive weight function as an estimate of the microcanonical entropy $\ln W^{(n+1)}(E) \approx -S(E)$ in an analog self-consistent approach. This microcanonical entropy is a good base for the discussion of the recursive updates scheme. One may now consider the change in entropy $\Delta S(E) = S(E + \Delta E) - S(E)$ and thus relate two energy values. Inserting the definition for $S(E)$ and the estimation of $\Omega(E)$ one obtains

$$\Delta S(E) = \ln\left(\frac{H(E+\Delta E)}{W(E+\Delta E)}\frac{W(E)}{H(E)}\right) \tag{5.27}$$

In the recursive scheme, this $\Delta S^{(n+1)}(E)$ is estimated using the weighted average of the previ-

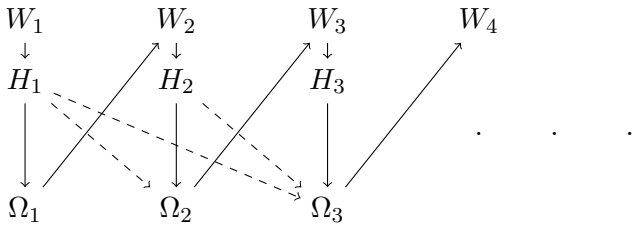

Figure 5.2.: Idea of the multicanonical recursion.

ous estimates $\Delta S^{(n)}(E)$ (including all previous iterations) and the current estimate Eq. (5.27), see Fig. 5.2. The quality of the estimate is weighted with the parameter $p(E)$. This parameter compares the values of the contributing histogram entries: if both entries have approximately the same value, the parameter has to be large, while the parameter needs to be small if one histogram entry is much larger than the other.

$$p^{(n)}(E) = \frac{H(E+\Delta E)H(E)}{H(E+\Delta E)+H(E)} \tag{5.28}$$

The accumulated parameter $p_{\mathrm{acc}}(E) = \sum_i^{n-1} p^{(i)}(E)$ is the collected weight of the estimate $\Delta S^{(n)}(E)$. The consecutive entropy change is thus

$$\Delta S^{(n+1)}(E) = \frac{p_{\mathrm{acc}}(E)\Delta S^{(n)} + p^{(n)}(E)\Delta S}{p_{\mathrm{acc}}(E) + p^{(n)}(E)}, \tag{5.29}$$

where $\Delta S^{(n)}(E) = \ln W^{(n)}(E) - \ln W^{(n)}(E+\Delta E)$ and $\Delta S(E)$ is given by Eq. (5.27). The new (logarithmic) weight function is obtained by the sum of entropy changes $\ln W^{(n+1)}(E) = -\sum_{E'<E} \Delta S^{(n+1)}(E')$. In the literature one often finds another formulation defining a weight relation $R(E) = W(E+\Delta E)/W(E)$. In this case most of the summands in the exponential of the weight functions cancel, which leads to $R^{(n+1)}(E) = \exp\left\{-\Delta S^{(n+1)}(E)\right\}$. Moreover, inserting Eq. (5.29) with all definitions and sorting according to weights and histograms, this yields

$$R^{(n+1)}(E) = R^{(n)}(E)\left[\frac{H^{(n)}(E)}{H^{(n)}(E+\Delta E)}\right]^{\kappa}, \tag{5.30}$$

where $\kappa = \frac{p^{(n)}(E)}{p_{\mathrm{acc}}(E)+p^{(n)}(E)}$. This is a more compact representation but somehow hides the insightful relation to the microcanonical entropy. The numerical implementation benefits from storing and working with logarithmic numbers, which requires logarithmic addition: $\ln C = \ln(A+B) = \max(\ln A, \ln B) + \ln\{1+\exp(-|\ln A - \ln B|)\}$ [122]. As mentioned before, this is due to numerical reasons when handling numbers that span several orders of magnitude: the finite number of decimal places on a computer causes problems, e.g., adding to a small number a very large one, followed by subtracting the same large number may not yield the same small number again.

One remaining point is the convergence criterion, or a combination of criteria, that determines when the weight iteration is finished and the production run begins. First, it is important

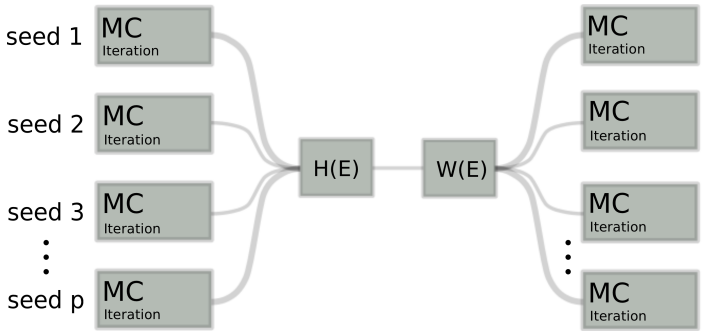

Figure 5.3.: Scheme of the parallel multicanonical method taken from Ref. [120]. Independent Monte Carlo simulations sample the same distribution, after a fixed number of updates their histograms are merged, the consecutive weights are determined and distributed to the parallel threads which continue with the next iteration.

that the full energy range is covered with sufficient entries in the histogram. I generally considered this to be valid if no histogram entry deviated from the average by more than a factor of 2. Additionally, I required the sampling processes to achieve sufficiently many tunnel events, passing from the high-energy bound to the low-energy bound or vice versa.

5.5. Parallel MUCA

Due to the large computational demand of condensation and aggregation phenomena especially for off-lattice systems and in the limit of large systems, it was unavoidable to optimize the simulation method. Since computer performance mainly increases in terms of parallel processing on multi-core architectures, we developed a parallel implementation of the multicanonical method [120]. The results of this section are published already in Refs. [17, 120, 128], where we tested the scaling properties for the Ising and q-state Potts model [120, 128], for a single bead-spring polymer [128] and for lattice gas condensation [17]. In all cases, we observed ideal scaling under suitable conditions.

The idea is to distribute the time consuming generation of statistics on p independent processes, a scheme is presented in Fig. 5.3. In each iteration, the processes perform equilibrium Monte Carlo simulations with identical weights $W^{(n)}(E)$ and measure individual histograms $H_i(E)$. After a chosen number of sweeps per iteration N_{iter}, these histograms are merged to a total histogram $H(E) = \sum H_i(E)$. The total histogram represents the sampled distribution of the fixed weight function with independent, uncorrelated contributions. It has at least the same quality as a histogram from a single-process simulation with the same amount of statistics. Now, the weight modification of choice may be applied with $H(E)$ in order to obtain $W^{(n+1)}(E)$, which is again distributed to all processes and the next iteration starts.

5.5.1. Performance for spin systems

I investigated the scaling properties of the parallel multicanonical method for the case of the two-dimensional Ising and q-state Potts model on a square lattice in Refs. [120, 128]. The Ising model exhibits a temperature driven second-order phase transition and the q-state Potts model (2D) shows a temperature driven first-order phase transition for $q \geq 5$ and a second-

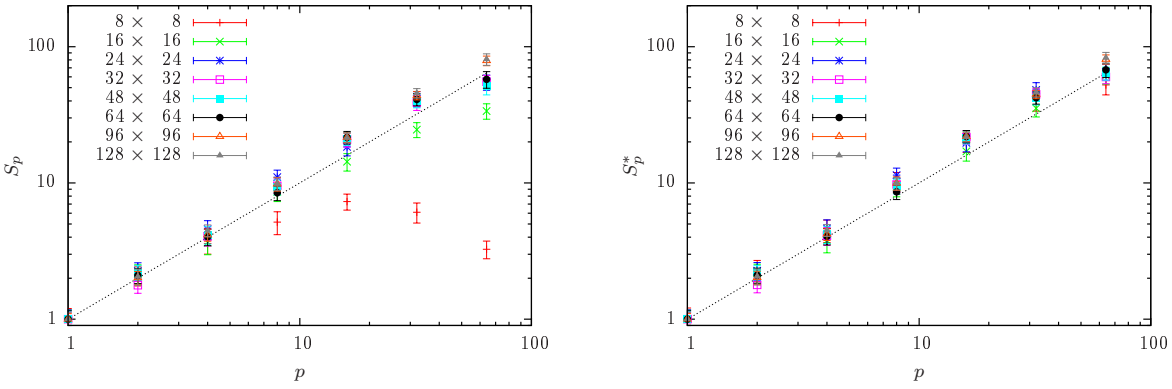

Figure 5.4.: Performance of the parallel multicanonical method for the Ising model with increasing system size: (left) the speedup factor and (right) the time-independent statistical speedup factor. The dashed line represents a perfect scaling $S_p = p$.

order phase transition otherwise. The Ising model is a widely applied model in statistical physics and a standard model for testing algorithms and scaling predictions. It is described by the Hamiltonian $\mathcal{H}^{(\text{Ising})} = -J \sum_{\langle i,j \rangle} s_i s_j$, with nearest-neighbor interaction, the spin variable $s \in \{-1, 1\}$ and the coupling constant J. This is equivalent to the $q = 2$ state Potts model, which is a generalized spin model with $\mathcal{H}^{(\text{Potts})} = -J \sum_{\langle i,j \rangle} \delta(s_i, s_j)$, where $q \in \{0, \ldots, q-1\}$ and $\delta(s_i, s_j)$ is the Kronecker-Delta function. In this section, I will consider $q \in \{2, 3, 4, 6, 8\}$.

For a fair comparison, I consider the optimal number of sweeps per core M_{opt} for each degree of parallelization p. This will influence the required number of iterations until convergence N_{iter}. A detailed discussion is given in Ref. [120], where the minimum of $N_{\text{iter}} M p$ vs. M is determined and the following dependence on linear system size L and degree of parallelization is obtained:

$$\begin{aligned} M_{\text{opt}}^{(\text{Ising})}(L,p) &= 5.7(5) \times L^{2+0.51(4)} \frac{1}{p} \\ M_{\text{opt}}^{(8\text{Potts})}(L,p) &= 24(4) \times L^{2+0.67(6)} \frac{1}{p}. \end{aligned} \tag{5.31}$$

This power-law behavior is characteristic for the explicit implementation, but is consistent with the scaling of multicanonical tunneling times in Refs. [117, 118, 129] and reflects that a random walk through energy space has to depend on the system size and the number of spin states q. Interesting to notice is the prefactor ratio which corresponds to the increase in q, from the $q = 2$ (Ising) to the $q = 8$ Potts model. I will make use of this, when comparing the scaling properties depending on q, interpolating Eq. (5.31) for the intermediate values. For a systematic study, the systems are thermalized only in the beginning and not in between iterations in order to avoid additional parameters.

Moreover, I apply two definitions for the speedup with p cores compared to single-core simulation [120]: the ratio of real simulation time t until convergence of the MUCA weights

$$S_p = \frac{t_1}{t_p}, \tag{5.32}$$

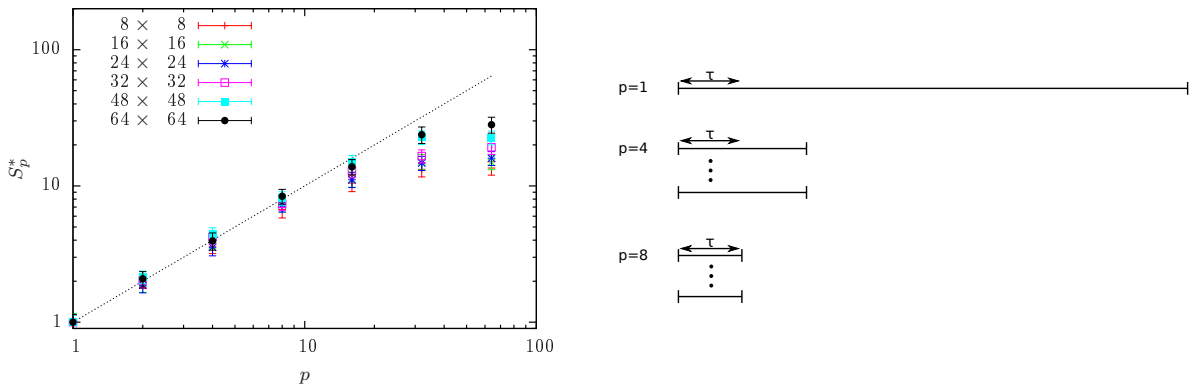

Figure 5.5.: (left) Statistical speedup factor for the 8-state Potts model with increasing system size. (right) Scheme of N_{iter}/core compared to the integrated autocorrelation time τ demonstrating the natural limit of parallelization [120].

and the time-independent statistical speedup, which is defined in terms of the total number of sweeps on each core until convergence $\bar{N}_{\text{iter}} M_{\text{opt}}(L,p)$,

$$S_p^* = \frac{[\bar{N}_{\text{iter}} M_{\text{opt}}(L,1)]_1}{[\bar{N}_{\text{iter}} M_{\text{opt}}(L,p)]_p}. \tag{5.33}$$

The subscript indicates the number of cores used. In both cases, error bars are obtained by averaging over 32 independent parallel-multicanonical simulations for each degree of parallelization. All simulations were performed on the compute cluster of our institute. If one considers only the distribution of statistics, the optimal scenario would be a linear speedup with slope one, $S_p = p$.

Figure 5.4 shows the results for the Ising model over several system sizes. It can clearly be seen that the parallel implementation leads to an ideal linear speedup up to 64 cores already for system sizes $L \geq 24$ in either definition. The statistical speedup is optimal for all considered system sizes. The "breakdown" for small system sizes in the real-time speedup (Fig. 5.4 (left)) may be easily explained by the extremely short simulation times within milliseconds which are difficult to measure on our compute cluster. While communication is kept to a minimum, only occurring in between iterations, this becomes relevant for small system sizes, when the iterations become very short. This is why the statistical speedup (Fig. 5.4 (right)) is a less biased observable, especially when communication infrastructure may vary drastically between the local compute cluster at our institute and international high-performance clusters. For the Ising model, we can see that for intermediate degrees of parallelization there actually occur speedups systematically larger than the optimal scaling. This is only possible, since the parallelization adds independent Markov chains with every degree of parallelization, which allow for a qualitatively better sampling of the underlying distribution than a single Markov chain. As a consequence, the estimate of the consecutive weights is improved and this may lead to a faster convergence. Thus, the parallelization also profits from a more systematic convergence to the desired flat distribution, which will be subject to further discussion below in Sec. 5.5.2.

The 8-state Potts model exhibits a first-order phase transition which is accompanied by

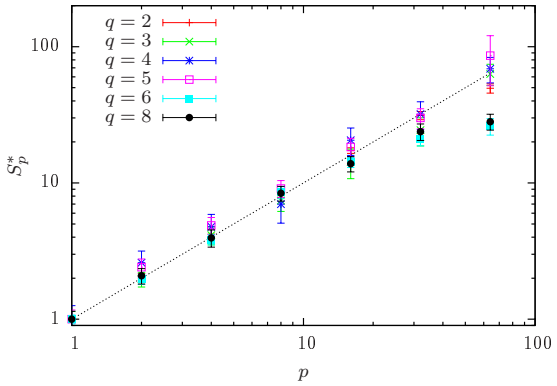

Figure 5.6.: Statistical speedup for different q-state Potts models ($q \leq 4$ second-order and $q > 4$ first-order phase transition) on a 64×64 lattice.

emerging barriers. The results of the performance study for the 8-state Potts model is shown in Fig. 5.5 (left) with increasing system size. The statistical speedup shows a saturation for large degrees of parallelization, where the final speedup increases with system size. This may be explained by the emerging barriers, which result in an increase of integrated autocorrelation time τ. If now, with increasing parallelization, the number of sweeps per core gets reduced to the order of τ, then this reduces the chance to efficiently cross the apparent barriers, see Fig. 5.5 (right). This was also observed for multimagnetic simulations [120, 128] where "hidden barriers" are known to emerge [15, 16, 48, 130].

As a cross-check of these arguments, Fig. 5.6 shows the scaling properties in dependence on the number of spin states q for a 64×64 square lattice. It can be seen, that for small q the scaling is ideal, while for large q the observed saturation occurs. This is consistent with the order of the temperature driven phase transition which is in two-dimensions second-order for $q \leq 4$ and first-order for $q > 4$. The case $q = 5$ may be considered as an exception because it is only weakly first-order. Thus, the emerging barriers and the involved integrated autocorrelation time give a natural limit to the degree of parallelization.

5.5.2. Performance for particle condensation and polymer aggregation

In order to justify the application of the parallel multicanonical method throughout my thesis, I tested the scaling properties for the cases of particle gas condensation [17] and polymer aggregation. The considered systems are the lattice particle gas in two and three dimensions as defined in Sec. 2.1, as well as the flexible bead-spring polymer system as defined in Sec. 3.1.

Figure 5.7 shows the performance of parallel multicanonical simulations for two- and three-dimensional lattice gas. In each dimension (d) we considered 1000 particles in a roughly constant volume. Other than for the Ising/Potts model, I did not predetermine the optimal number of updates per sweep and number of cores but assumed the same p-dependence as before, namely $25600/p$ sweeps in both cases. A single sweep consists of L^d updates. This choice was motivated by initial parallel multicanonical simulations that consistently converged and should serve as a day-to-day application example. For the statistical averages we consider 32 simulations for each degree of parallelization. The average number of iterations in Fig. 5.7 (left) shows a clear decrease with increasing number of cores. This shows the previously mentioned

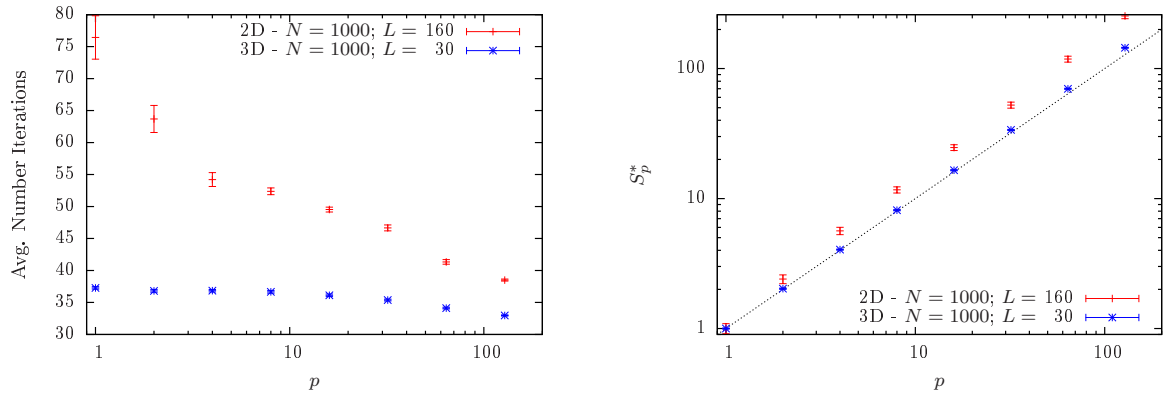

Figure 5.7.: Performance of the parallel multicanonical method for lattice gas condensation in two and three dimensions: the average number of iterations (left) shows that an increase in the number of independent Markov Chains improves the convergence and the statistical speedup (right) shows an (even better than) ideal behavior. The chosen parameter were not optimized as in the case of the Ising/Potts model.

improvement of sampling with increasing number of independent Markov Chains. The convergence relies on some chosen bounds and criteria and the number of iterations until these are reached vary depending on the initial seed of the random number generator. For a single process, the Markov Chain is correlated and with it the local contributions to a given distribution. This situation changes when several independent Markov Chains contribute to the sampling of the same distribution. The quality is improved and the required number of iterations until convergence may get reduced, as long as the individual Markov Chains are larger than the integrated autocorrelation time, see Sec. 5.5.1. However, this also shows that especially for the two-dimensional case the choice of parameters was not optimal.

This is also reflected in the statistical speedup, Fig. 5.7 (right), which shows an (better than)

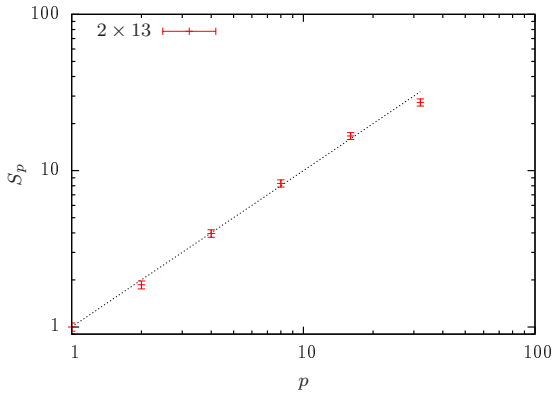

Figure 5.8.: Simulation-time speedup for the aggregation of two polymers of length $N = 13$. The speedup shows an ideal behavior that may be expected to continue also for more than 32 cores.

ideal speedup. While the two-dimensional case clearly seems to be a special case and depends on the chosen parameter, the general picture of an ideal speedup may be supported for the case of lattice gas condensation. The situation also remains unchanged for flexible off-lattice bead-spring polymers. For single polymer collapse, we showed an ideal speedup ($N = 40, 80$) for up to 32 cores [128]. For aggregation of two flexible polymers ($N = 13$) Fig. 5.8 shows the simulation-time speedup that again presents ideal scaling. There is no reason to expect a different result for semiflexible polymer aggregation or the aggregation of a lattice polymer (iSAW) model. The presented results justify to apply the parallel multicanonical method for the considered models throughout this thesis.

5.6. Adaptive update ranges for off-lattice updates

For updates with non-equal selection probability (p^s in Sec.5.1), there exists an efficient Metropolis-Hastings like scheme to tune the acceptance rate without breaking detailed balance, which is also applicable in combination with the multicanonical method: adaptive update ranges [70]. I will briefly recapture the proposed method and the general formulation and refer to Ref. [70] for additional details. The explicit update-dependent realizations will be given in Sec. 5.7, where the updates are introduced.

The main goal is to choose energy-dependent update ranges in order to obtain a sufficiently large acceptance rate especially at low energies and correct for the emerging detailed balance violation. Why may detailed balance be violated? Recall that the detailed balance condition Eq. (5.7) leads to the general acceptance probability Eq. (5.8) with the probability distribution $P(q)$ and the selection probability $p^s(q' \to q)$. Considering the above described multicanonical ensemble, the probability distribution becomes $W(E)$ and the acceptance probability to go from a state of energy E to a state of energy E' is again

$$p^a(E \to E') = \min\left(1, \frac{W(E')p^s(E' \to E)}{W(E)p^s(E \to E')}\right), \tag{5.34}$$

where now $p^s(E \to E') \neq p^s(E' \to E)$ in general. If this inequality is not taken into account, then the acceptance probability is biased and this leads to detailed balance violation and errors in the sampled distribution. There are, in principle, two possible scenarios now that need to be considered. In the best case, the selection probabilities are different but not zero and the occurring violation of detailed balance may be corrected by choosing an appropriate acceptance weight according to Eq. (5.34). If, however, the selection probability for the reversed update is zero, then this is a "hard detailed balance violation" and the update needs to be rejected!

This may be illustrated on the example of a uniform single particle displacement within a box of length L around its current position $\boldsymbol{R}$. The proposed new position is $\boldsymbol{R'} = \boldsymbol{R} + \boldsymbol{r}$, where $\boldsymbol{r} = (x_1, x_2, x_3)$ is a random vector with $|x_i| \leq L/2$. The selection probability for this update is given by the inverse volume of the box, $p^s(E \to E') = L^{-3}$, which may be seen as a proposal density. Now, imagine that the update range for position $\boldsymbol{R'}$ with energy E' is L'. If $L' > L$ there will be no problem because $\boldsymbol{R} = \boldsymbol{R'} - \boldsymbol{r}$ may be constructed with $|x_i| \leq L'/2$. However, the selection probability changed, namely $p^s(E' \to E) = L'^{-3} < L^{-3}$. Thus, the probability to go from state $\boldsymbol{R}$ to $\boldsymbol{R'}$ is larger than the reverse scenario and this is corrected in the initial acceptance ratio via Eq. (5.34). The situation changes if $L' < L$; it might occur that the reverse transformation $\boldsymbol{R} = \boldsymbol{R'} - \boldsymbol{r}$ is no longer allowed since $|x_i| > L'/2$. Then, the proposed update $E \to E'$ will have to be rejected (because there is no legal "back transformation") in order to

avoid hard detailed balance violation. For the displacement in a box, we may summarize

$$p^a(E \to E') = \begin{cases} \min\left(1, \frac{W(E')L^3}{W(E)L'^3}\right), & \text{if } |x_i| \leq L' \\ 0 & \text{else.} \end{cases} \tag{5.35}$$

For all updates I consider, the scenario will be similar: a selection probability is derived from the possible update volume that depends on a single parameter, the *update range*. The update range is adjustable as a function of the energy and for now denoted $U(E)$. This function starts out as a constant function and is tuned within the multicanonical iteration. After each update proposal $E \to E$' (accepted or rejected), the update range of the initial energy $U(E)$ is modified according to the rule [70]

$$U^{\text{new}}(E) = \begin{cases} (1+2\epsilon)\ U(E) & \text{if } E > E' \\ (1-\ \ \epsilon)\ U(E) & \text{if } E \leq E', \end{cases} \tag{5.36}$$

where $0 < \epsilon \ll 1$. This should lead to an average acceptance of about $60 - 66\%$ [70], which I can confirm. Since I employ parallel multicanonical simulations, I tune the update ranges on every processor and average the results after every iteration. In addition, roughly every 10 000 updates, I locate the lowest energy sampled so far E_{low} and the energy with the smallest update range $U(E_{\text{min}})$, where $E_{\text{min}} > E_{\text{low}}$, and adjust all $U(E)$ for $E < E_{\text{min}}$ with the average $\overline{U} = \sum_{E_{\text{low}}}^{E_{\text{min}}} U(E)$. For less noise on the update range function, I further smooth the full range with a low-pass filter, essentially averaging over a small neighborhood of each function value.

It has to be noted that especially in the initial multicanonical iterations there may occur large fluctuations on the boundary entries of the histograms, where the estimate of the weights rely on only little statistics from the previous iteration. This may lead to severe problems in the tuning procedure but can be circumvented by introducing a lower energy bound for the tuning range $E_{\text{t,low}}$ below which the update range is not modified. This energy bound may be chosen as the lowest energy reached in the second to last iteration, which is then usually already well sampled in the current iteration. Also, since the update range constantly changes in the tuning process, this actually violates detailed balance in the iteration procedure. I usually choose a tuning parameter $\epsilon = 10^{-4}$, which leads to only small violations and does not disturb the usual multicanonical convergence. However, for the final production run, the update range function has to be fixed just like the weight function in order to fulfill detailed balance and with it obtain reliable equilibrium estimates.

5.7. Lattice and off-lattice updates

In this section, I will briefly explain the Monte Carlo updates used from lattice particle condensation to off-lattice polymer aggregation. Instead of discussing the implementation in great detail, I will focus on the concepts and difficulties. Modifications necessary for the adaptive update ranges in Sec. 5.6 are provided where possible. It is my experience that the choice of updates is crucial for an efficient, ergodic outcome of any applied method to a given problem. An exemplary demonstration is presented in Sec. 5.7.6. Consistency checks with exact or independent results are provided in the corresponding previous sections.

In most cases, I chose a combination of updates to allow efficient sampling of both the high- and low-energy domains. In this case, I randomly chose from the set of updates with different,

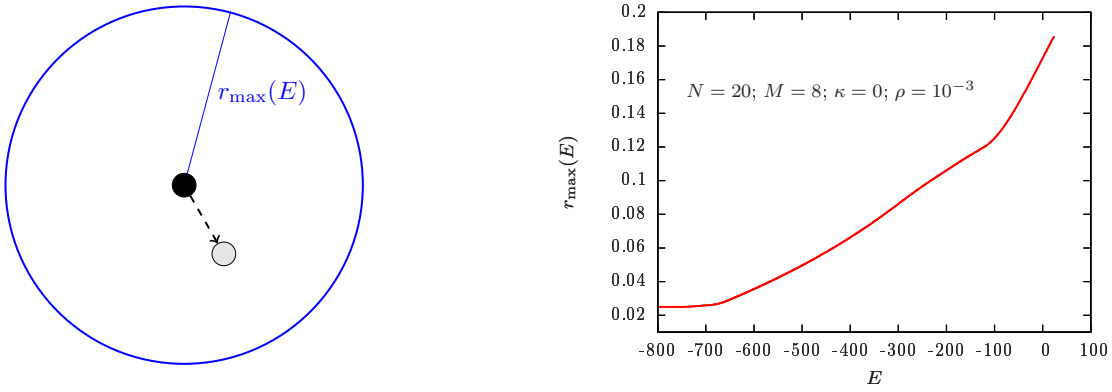

Figure 5.9.: (left) Schematic presentation of the single-monomer displacement move within a sphere of size $r_{\max}$. (right) Result of the adaptive update ranges on the example of eight flexible 20mers.

heuristic probabilities in order to ensure both ergodicity and good sampling.

5.7.1. Single monomer/particle displacement

The basic update for off-lattice bead-spring polymers, off-lattice particle gas and lattice gas models is the single monomer or particle displacement. This update allows in principle to reach any conformation, while for polymer aggregation the combination with additional moves is recommendable (see below). For off-lattice systems, the single particle is shifted to a uniformly distributed random position $\boldsymbol{R}'$ within a sphere with radius $r_{\max}$ around its current location $\boldsymbol{R}$ (see Fig. 5.9 (left)). In case of lattice gas, I only considered displacements to either a random nearest neighbor or a jump to any random lattice site (equivalent to a Kawasaki update scheme).

For off-lattice particles displaced in a sphere of radius $r_{\max}$ we may apply the adaptive update ranges from Sec. 5.6. Then, the maximal update range becomes a function of the energy $r_{\max}(E)$ and the selection probability (proposal density) is give by the inverse volume $p^s(E \to E') = \left(\frac{4}{3}\pi r_{\max}(E)^3\right)^{-1}$. This yields for a proposed single shift by $\boldsymbol{r}$ the acceptance probability

$$p^a(E \to E') = \begin{cases} \min\left(1, \frac{W(E')r_{\max}(E)^3}{W(E)r_{\max}(E')^3}\right), & \text{if } |\boldsymbol{r}| \le r_{\max}(E') \\ 0 & \text{else.} \end{cases} \tag{5.37}$$

An example of a tuned maximal shift range is presented in Fig. 5.9 (right) for eight flexible polymers of length $N = 20$.

5.7.2. Bond rotation and pivot

Another quite generic update for off-lattice polymers is the single-bond rotation update, or spherical update [74, 131]. Here, a random bond is defined as the zenith direction and rotated along a spherical cap with opening angle $2\theta_{\max}$. This is achieved by uniformly choosing the polar angle θ and azimuthal angle φ in the ranges $\cos\theta_{\max} \le \cos\theta \le 1$ and $0 \le \varphi < 2\pi$, respectively. This update is bond-length conserving up to numerical precision. The remaining

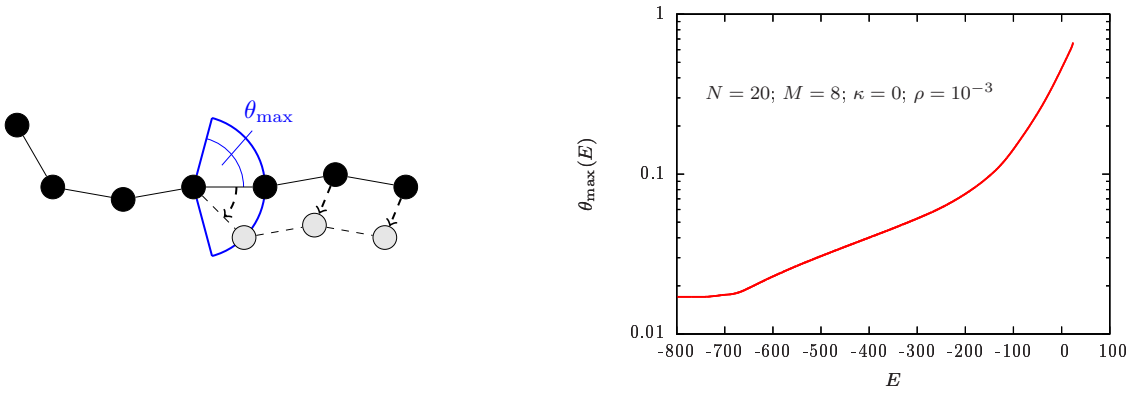

Figure 5.10.: (left) Schematic presentation of the bond-rotation move within a cone of size $2\theta_{\max}$. (right) Result of the adaptive update ranges on the example of eight flexible 20mers.

"end" of the polymer is simply translated by the resulting shift, see Fig. 5.10 (left) for a two-dimensional sketch. In practice, a random monomer i is chosen together with a random polymer end (either direction along the polymer starting from monomer i).

Again, this update is suitable for the application of adaptive update ranges (Sec. 5.6). Here, the opening angle of the cap becomes a function of the energy, $\theta_{\max}(E)$. The inverse surface of the cap (a simple integration in the before mentioned range) determines the selection probability $p^s(E \to E') = (2\pi\,(1 - \cos\theta_{\max}(E)))^{-1}$. The acceptance probability is depending on $\cos\theta_{\max}$ because $\cos\theta$ needs to be uniformly distributed, i.e.,

$$p^a(E \to E') = \begin{cases} \min\left(1, \frac{W(E')(1-\cos\theta_{\max}(E))}{W(E)(1-\cos\theta_{\max}(E'))}\right), & \text{if } \theta \le \theta_{\max}(E') \\ 0 & \text{else.} \end{cases} \tag{5.38}$$

An example of a tuned maximal polar angle $\theta_{\max}$ is presented in Fig. 5.9 (right) for eight flexible polymers of length $N = 20$.

For lattice polymers, I applied the *pivot* update [67, 132, 133] instead, which allows for discrete rotations of a full polymer end about a random monomer (the pivot point). The natural restriction of the opening angle is then 90 degrees. For more than one polymer, again one of both ends needs to be selected at random.

5.7.3. Slithering snake

Throughout my thesis, I additionally considered the slithering snake update for lattice polymers, while it is also applicable to off-lattice polymers in principle. Here, a random end of the polymer is chosen at which a new randomly oriented bond is created. For lattice polymers this means to draw a random next-neighboring site to the last monomer. The last monomer on the other end is removed. This is accepted with the usual acceptance probability considering all hard constraints (e.g. self-avoidance, walls, etc.). It is important that the polymer end is randomly chosen because otherwise systematic errors may be (and have been) encountered, especially if an unfavorable sequence of updates with fixed order is chosen.

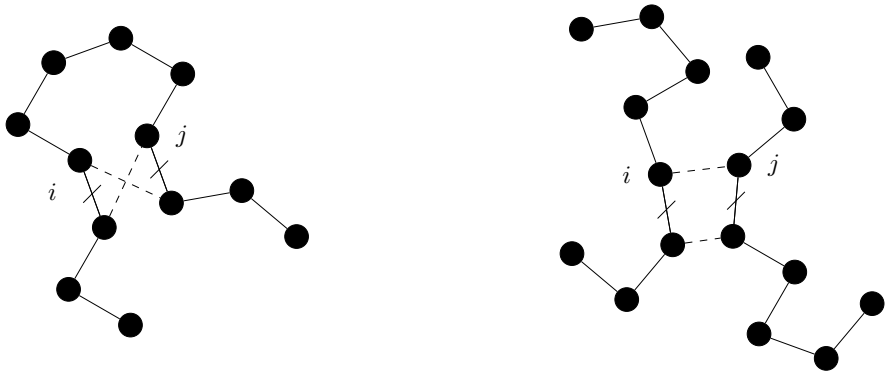

Figure 5.11.: Schematic presentation of the intra-polymer bridging (left) and inter-polymer bridging (right) move. The current implementation always considers double bridges.

5.7.4. Polymer translation

When several polymers are considered at low density, it is advantageous to additionally use full polymer translation updates. This is comparable to the particle displacement in the particle condensation and enhances the sampling of entropy. As in Sec. 5.7.1, a random displacement vector in a sphere of size $r_{\max}$ is chosen and all monomers belonging to the selected polymer are shifted by this vector. In principle, this would also allow to apply adaptive update ranges, however, I purposely did not use them in order to allow large translations out of the aggregate even at low energy.

5.7.5. Double-bridging

An important Monte Carlo update that I considered throughout my studies is the double-bridging or double-pivot move [134–136]. This update was crucial in order to "easily" reach and evenly sample low-energy (amorphous) aggregated phases. In exemplary test cases with several flexible polymers, this update made the difference of a successful simulation in reasonable time or not.

In order to ensure detailed balance, the implementation of the double-bridging update move that I applied always acts on pairs of bonded monomers. This is different to Ref. [134], where re-bridging of trimers is considered, and it makes use of the polymer model property that the average bond-length r_0 is of the order of the monomer extension σ [137]. In addition, the constraint of monodisperse homopolymer chains needs to be fulfilled which leads to a reduced set of possible re-bridging moves. First, a random monomer i of the system in state q is selected. Then all monomers within a radius R_{DB} are stored as possible bridging partners $N_{DB}(i,q)$, including the previously bonded ones. If the current bond length would be larger than R_{DB}, then the update is rejected. Using the FENE potential, Eq. (3.4), a suitable choice is the maximal extension $R_{DB} = r_0 + R$. Among the bridging partners, a random bridging partner monomer j is selected, which is either within the same polymer (intra-polymer bridging) or in another polymer (inter-polymer bridging), see also Fig. 5.11. In both cases, the selection probability for a bridging partner is then $p^s(i, q \to q') = 1/N_{DB}(i,q)$. Note that N_{DB} is always

non-zero because each monomer has at least one bonded bridging partner by construction. The same conformation q' can be obtained via several monomers (all involved monomers in the re-bridging). Since the double bridge update does not alter the monomer positions $N_{DB}(i,q) = N_{DB}(i,q')$. While the current implementation may not be the most efficient, it ensures that this remains true for all involved bridging monomers m such that the selection probability of the forward and reverse update is the same $p^s(q \to q') = \frac{1}{NM}\sum p^s(i, q \to q') = p^s(q' \to q)$ and detailed balance is not violated. However, since the bonds are extensible it needs to be checked in addition that all old and new bonds are shorter than R_{DB} because otherwise this would result in a hard detailed balance violation.

In case of the intra-polymer double bridging in Fig. 5.11 (left), the new bond $(i) - (j)$ requires a suitable second bond either between $(i+1)-(j+1)$ or $(i-1)-(j-1)$. This ensures monodisperse polymers and avoids a separation of a single polymer into a closed loop and an additional piece. Also, if i is at one end of the polymer this does not result in the so called backbite move but always remains an update that rewires two bonds. Thus, the problems with detailed balance discussed for the backbite update in Ref. [138] are avoided by this definition because choosing the same monomer i again would yield the same update probability in both forward and reverse direction. In my implementation for off-lattice polymers, the monomers enclosed by the pair i, j are merely exchanged. For lattice polymers, I did not consider the intra-polymer case.

For the inter-polymer double bridging update in Fig. 5.11 (right), there are three possible cases. Next to the new bond $(i) - (j)$ there may now be either the bond $(i+1) - (j+1)$, $(i-1)-(j-1)$ and $(i \pm 1)-(j \mp 1)$. In all cases, the condition of monodisperse polymers needs to be verified in addition. My implementation for off-lattice polymers follows this description, again exchanging corresponding pairs of monomers. For lattice polymers, my implementation is following Ref. [135].

5.7.6. How to choose a proper combination of updates

As mentioned before, the choice of Monte Carlo updates may make the difference of a successful sampling or not. Using the multicanonical method with non-efficient moves may lead to slow convergence, a reduced energy range, and few tunnel events. Tunnel events are here defined as the number of times the system evolves from the high-energy to the low-energy range or vice versa.

Using more than one update move may be realized by choosing a fixed set of updates that are repeated over and over again. However, I have encountered small detailed balance violations using an inconvenient set of updates including an asymmetric slithering snake update move, treating both ends separately but including both versions in the update set. As a consequence, it is advantageous to use symmetric implementations (including e.g. both polymer ends, see also the bond-rotation and slithering-snake update) if possible. Moreover, in order to choose arbitrary fractions of occurrence I generally choose a random update from a set of n updates with given probabilities p_i, $\sum_1^n p_i = 1$. In practice, a random number r is compared to $\sum_1^{i'} p_i$ and the largest possible i' for which $\sum_1^{i'} p_i < r$ determines the update move.

In order to test and demonstrate the efficiency of the chosen set of updates, I performed a test simulation with $M = 8$ polymers of length $N = 13$ and density $\rho = 5 \times 10^{-3}$, which is identical up to the probabilities p_i in the sweeps ($M \times N$ updates). The simulation uses 32 cores and starts with a short parallel tempering simulation (3 million sweeps in the temperature range $[0.5, 2]$), followed by a parallel multicanonical simulation with trivial weight updates

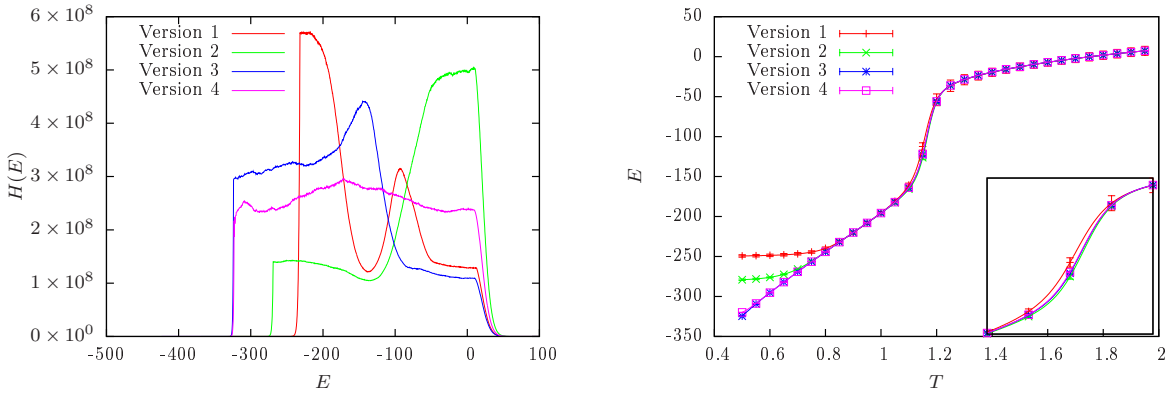

Figure 5.12.: Comparison of different combinations of updates from Table 5.1 for a 8×13 flexible polymer system. The final histogram (left) shows the covered range and gives a hint on the weak sampling, while the canonical energy (right) clearly shows the drawback of purely local moves.

(see Sec. 5.4.2) in the energy range determined by the parallel tempering part. Afterwards a production run with 2.56×10^9 sweeps is performed. The set of updates includes single-monomer displacement (S), polymer translation (T), bond rotation (R) and double-bridging (B) moves. Also, adaptive update ranges (see Sec. 5.6) are applied for (S) and (R) within the convergence part but no longer in the production run. Details are provided in Table 5.1 including the update selection probability p_i, the lower-energy-bound from the parallel tempering simulation $E_{\text{min,PT}}$, the number of iterations until convergence, the simulation time, and the number of tunnel events in the final production run.

Figure 5.12 (left) shows the histograms from the final production run. The covered energy range is a result from the parallel tempering run, while the fluctuations on the histogram show the "quality" of the random walk in energy. The reweighted estimate of the canonical expectation value of the total energy E is shown in Fig. 5.12 (right). It can clearly be seen that using only local single-monomer displacements (version 1) results in a reduced energy range that is moreover weakly sampled and does not allow to reweight down to $T < 0.9$. This may be explained by the large correlation times due to only small changes of conformations. Adding polymer translations (version 2) enhances the sampling of the high-energy phase (where the polymers are separated) and also brings the polymers together more often, resulting in a slightly

Table 5.1.: Comparison of different combinations of updates for a 8×13 flexible polymer system with $\rho = 5 \times 10^{-3}$. In each step the updates were randomly chosen from a set of updates including single-monomer displacement (S), translation (T), bond-rotation (R) and double-bridge (B) moves.

Version	Updates	p_i	$E_{\text{min,PT}}$	N_{iter}	time(h)	Tunnel-events
1	S	1.0	-232.00	90	10.6	73
2	S, T	0.8/0.2	-269.00	57	17.5	118
3	S, T, R	0.5/0.1/0.4	-324.00	38	25.6	300
4	S, T, R, B	0.5/0.1/0.2/0.2	-323.00	15	19.2	12022

lower energy and increased number of tunnel events, see also Table 5.1. This is also reflected in the total energy in the vicinity of the aggregation transition (see inset in Fig. 5.12 (right)), where the expectation value of version 1 coincides within the error with the remaining version but clearly deviates. The reason for this is the increased sampling of the aggregation transition associated with more tunnel events. Applying non-local updates, namely the bond-rotation update in version 3 drastically lowers the energy bound and allows to reweight down to the desired temperature $T_{\min} = 0.5$. Using in addition the double-bridging move (version 4) results in a flatter histogram after fewer iterations with a large increase in tunnel events. This is due to the fact that the low-energy state of flexible polymers is an amorphous aggregate (see Sec. 4.2 and Ref. [95]) where the double-bridging update allows major conformational changes. The canonical expectation values are almost indistinguishable between version 3 and 4.

This suggests that the implementation of the involved moves satisfies detailed balance. Moreover, it emphasizes the need for an appropriate combination of proper updates suitable to the problem at hand for a successful and efficient (multicanonical) simulation. For (flexible) polymer aggregation, version 4 is a good choice.

5.8. Error estimation

In order to judge the quality of an estimator, it needs a proper error estimation. I mainly applied two methods to estimate the error. One possible measure of sampling quality is the number of tunnel events, i.e., how often a simulation passed from the high-energy bound to the low-energy bound or vice versa. If a production run is long enough and includes many tunnel events (or possibly several "independent" production runs as will be the case for the parallel multicanonical method, see Sec. 5.5), the *binning error analysis* provides a quick and decent error estimate. The time series is split into N_B equally sized bins. For each bin, the according estimator $O_{B,n}$ is calculated. Under the assumption that the estimators are uncorrelated, their variance may be obtained with the standard (unbiased) estimator, which leads to the squared binning error as a statistical error [6]

$$\epsilon^2_{B,O} = \frac{1}{N_B(N_B-1)} \sum_n \left(O_{B,n} - \overline{O}_B\right)^2, \tag{5.39}$$

where $\bar{O}_B$ is the average of the binned estimators. If the data is more correlated, however, there is a way to deal with this correlation using the *jackknife error analysis* [139]. Again, the time series is split into N_B blocks, but now the estimators $O_{J,n}$ are calculated over the full time series excluding the current bin. That way, the jackknife estimators are trivially correlated which may be corrected by multiplying $(N_B-1)^2$ and the squared error becomes [6]

$$\epsilon^2_{J,O} = \frac{N_B-1}{N_B} \sum_n \left(O_{J,n} - \overline{O}_J\right)^2, \tag{5.40}$$

where $\bar{O}_J$ is equivalently the average of jackknife estimators. While this of course takes more time, it is an error estimator without correlations.

Throughout this thesis, I always consider data points with error bars from time-series reweighting together with connected lines at higher precision from histogram reweighting. Using multicanonical simulations, reweighting procedures may only be applied as long as the canonical probability distribution is covered by the multicanonical histogram. With respect to

the canonical estimators, this limit of reweighting range may be noticed in unexpected behavior of the otherwise monotonic total energy (e.g. if the energy at low temperatures suddenly becomes constant, see also Sec. 5.7.6) and consequently in the specific heat. For continuous systems, the ground state is usually not sampled for non-zero temperatures (this may differ for discrete models). Thus, a constant energy accompanied by a vanishing heat capacity usually indicates the lower temperature bound and may introduce artificial transition peaks.

6. Summary and outlook

In this thesis, large-scale multicanonical simulations were combined with theoretical arguments in order to study phase (separation) transitions from particle condensation to semiflexible polymer aggregation. An additional focus was put on the structural motifs of polymer aggregates, where stiffness was identified as a key parameter. These studies required further development of advanced simulation techniques, which I discussed in detail. This included a parallel implementation of the multicanonical method with extensive performance tests on all relevant model systems discussed [17, 120, 128].

First-order condensation and aggregation transition. Particle condensation/evaporation is a standard example of a first-order phase transition in the limit of an infinite number of constituents. Theoretical considerations show that the transition separates a homogeneous gas phase and an inhomogeneous phase of a droplet surrounded by vapor. This leads to scaling predictions for canonical observables, the free-energy barrier, and microcanonical signatures, which were presented in Chap. 2. For (semiflexible) polymer aggregation (Chap. 3-4) the same qualitative behavior was observed for an increasing number of polymers of fixed length. Especially, the microcanonical analysis showed a prominent "backbending effect". The positive slope of the corresponding inflection points was shown to decrease with increasing polymer number. In this limit, polymer aggregation is thus a first-order phase transition in the common sense, separating an entropy-dominated and an energy-dominated phase. The presented results rely on finite-size-scaling analyses at a fixed density. For condensation and aggregation phenomena, this approach is orthogonal to the common fixed-temperature considerations. For the condensation transition temperature of a lattice gas, this orthogonal approach was shown to yield results consistent with exact solutions and low-temperature series expansions [21]. In addition, this approach highlights complementary aspects, with both advantages and disadvantages. One unexpected finding for polymers is an intermediate scaling behavior of the aggregation temperature and the transition rounding, which deviates from the large-system scaling expectations (Chap. 3-4 and Ref. [42]). This is, however, consistent with a mesoscopic scaling regime observed for particle condensation (Chap. 2 and Ref. [21]). In this regime, the aggregate/condensate includes a majority of the constituents (and thus dominates the inhomogeneous phase), which influences the interplay of fluctuation entropy and surface tension relevant for the scaling expectations. This is consistent with a relation to non-periodic first-order phase transitions between homogeneous phases, where the aggregate/condensate may be interpreted as the relevant embedded subsystem. Particle condensation and dilute polymer aggregation thus show strong similarities. This is not unexpected when polymer aggregation is interpreted as condensation of extended objects. In this case, dilute polymer aggregation is also expected to be a first-order phase transition between a homogeneous phase (solute) and an inhomogeneous phase (aggregate plus unattached polymers). However, the large-system scaling would be expected only for length scales much larger than the single polymer extension.

Density dependence of condensation and aggregation transition. At the homogeneous-to-inhomogeneous phase transition, energy minimization balances entropy maximization. Keeping the number of constituents constant and changing the density thus allows, in leading order, to study the entropic contribution in isolation. From ideal-gas considerations in the microcanonical ensemble, one can estimate that the transition temperature depends linearly on the logarithm of the density. This was confirmed for particle condensation with periodic boundary conditions (Chap. 2) and for semiflexible polymer aggregation in spherical confinement (Chap. 3-4 and Refs. [42, 65]). Rescaling the transition temperature with the finite-size-scaling results at fixed density yields a consistent data collapse in all considered cases. In addition, the spherical confinement allowed to safely investigate the onset of the dilute limit for polymer aggregation, excluding the possibility of a single polymer to self-interact across the boundaries. Comparing the single-chain and many-chain volume fraction yields an estimate of the minimal confinement size at which the system can be considered dilute. For the considered example, this estimate is in remarkable agreement with the numerical data. In the dilute regime, separated polymers above the aggregation temperature behave like isolated chains. The aggregation transition then dominates the single-polymer collapse for flexible polymers: starting in the high-temperature phase (above the aggregation temperature), single polymers begin to undergo the continuous collapse transition down to the point of aggregation, where the collapse process is then partially reverted.

Structural phases of semiflexible aggregates. For semiflexible Θ-polymers, stiffness plays a key role in determining structural motifs of aggregates [95], ranging from amorphous aggregates to polymer bundles. This was discussed in Chap. 4 for several small system sizes. Flexible polymers tend to form amorphous aggregates, spherical objects in which the polymers are more or less uncorrelated. By contrast, stiffer polymers form bundles, ordered structures in which polymers are aligned. At lower temperatures these bundles become twisted, in good agreement with wormlike-chain-based descriptions applicable in this limit of rather stiff Θ-polymers. The discussion was supported by an effective bundle parameterization. It was shown that in this case the energy is minimized by a twist for all stiffnesses. This has exciting implications, e.g., the possible formation of twisted flexible bundles under external force. In addition, semiflexible polymers feature an intermediate-stiffness regime, where both amorphous aggregates and polymer bundles are possible. A microcanonical analysis showed that the transition from amorphous aggregates to bundles becomes second-order like in this case. For the full stiffness range, the first-order aggregation transition is accompanied by a free-energy barrier. This barrier was shown to increase with stiffness and hence with structural order of the aggregate. Additional finite-size-scaling results for the free-energy barrier of semiflexible polymers support this result and suggest that it is a generic feature of the structural order to increase the free-energy barrier, which is in agreement with recent claims in the literature. The presented results thus expand the understanding of transition mechanisms leading to the formation of mesoscopic structures and the corresponding structural properties of “a few” polymers. This is important, because relevant systems are usually neither empty, nor much larger than the polymer extension. In this way, the study of (isolated) finite systems in equilibrium may help to understand specific transitions occurring in nature. In addition, the application of finite-size scaling may identify and support equivalences even without the formal definition of a thermodynamic limit.

Outlook: Of course, I was not able to answer every question that I raised within this thesis. This was partially due to the large computational demand. One promising new approach are "physical" microcanonical Monte Carlo simulations [58, 59] in combination with a replica exchange scheme. For technical reasons one could then also choose a polymer model in which bonded monomers are connected with usual harmonic springs instead of the FENE potential. Regarding the scientific content, an extended study of the intermediate scaling regime of semiflexible polymers would be interesting, in order to better investigate the effect of the aggregate shape on the scaling. Of special interest would be a more detailed scaling analysis of the free-energy barrier and a systematic investigation of the free-energy landscape. Extending the exact enumeration study of two lattice polymers by Monte Carlo simulations would bring further insight, generalizing the equivalence between polymer binding and adsorption known from directed polymers to undirected flexible polymers. A particularly promising new topic is the aforementioned possibility to construct twisted bundles of flexible polymers under external force. If achievable, this could lead to exciting applications, e.g., aperture devices on the molecular level.

A. Appendix

Table A.1.: Condensation of 2D lattice gas with fixed density $\rho = 10^{-2}$. The condensation temperature T_c and rounding ΔT are obtained from the specific heat peak and the fraction $\langle\tilde{\eta}\rangle_{T_c}$ of particles in the largest cluster is determined at T_c. The equal height temperature T_{eqh} is estimated from a binary-search histogram reweighting scheme and the free-energy barrier is estimated at this temperature. Error bars, if given, are obtained from jackknife error analysis.

L	N	T_c	ΔT	$\langle\tilde{\eta}\rangle_{T_c}$	T_{eqh}	$\beta_{eqh}\Delta F$
20	4	0.23427(15)	0.09481(7)	1.0(0)	-	-
30	9	0.27102(15)	0.09260(16)	0.81503(16)	-	-
40	16	0.29315(6)	0.07260(11)	0.72686(17)	-	-
50	25	0.30800(19)	0.05888(12)	0.66910(20)	-	-
60	36	0.31906(5)	0.0489(7)	0.61223(21)	-	-
70	49	0.32753(14)	0.0410(6)	0.56242(24)	-	-
80	64	0.33423(6)	0.03486(20)	0.51834(10)	-	-
90	81	0.33969(5)	0.02968(24)	0.48066(14)	-	-
100	100	0.344182(22)	0.02512(14)	0.45298(12)	-	-
120	144	0.351199(28)	0.01839(8)	0.39803(10)	0.352	-
140	196	0.35640(4)	0.01375(8)	0.35792(10)	0.357	-
160	256	0.36041(4)	0.010541(29)	0.32574(10)	0.361	-
180	324	0.363612(18)	0.00826(8)	0.30094(15)	0.364	0.410(21)
200	400	0.366229(22)	0.00665(4)	0.2766(18)	0.366	0.553(21)
250	625	0.371054(12)	0.004293(29)	0.24022(16)	0.371	1.022(29)
300	900	0.37439(4)	0.003015(15)	0.21371(12)	0.374	1.442(28)
400	1600	0.378777(22)	0.001778(6)	0.1785(8)	0.379	2.44(5)
500	2500	0.381605(25)	0.001205(6)	0.1564(8)	0.381	3.43(8)
600	3600	0.383582(13)	0.000889(6)	0.1402(7)	0.383	4.46(6)
700	4900	0.385081(10)	0.0006987(18)	0.1277(6)	0.385	5.21(5)
800	6400	0.38622(4)	0.000562(5)	0.1185(17)	0.386	6.43(10)
900	8100	0.387183(28)	0.000473(4)	0.1100(10)	0.387	7.15(10)
1000	10000	0.387988(15)	0.0004033(11)	0.1032(8)	0.388	7.83(10)

Table A.2.: Condensation of 3D lattice gas with fixed density $\rho = 10^{-2}$. The condensation temperature T_c and rounding ΔT are obtained from the specific heat peak and the fraction $\langle\tilde{\eta}\rangle_{T_c}$ of particles in the largest cluster is determined at T_c. The equal height temperature T_{eqh} is estimated from a binary-search histogram reweighting scheme and the free-energy barrier is estimated at this temperature. Error bars, if given, are obtained from jackknife error analysis.

L	N	T_c	ΔT	$\langle\tilde{\eta}\rangle_{T_c}$	T_{eqh}	$\beta_{eqh}\Delta F$
14	27	0.38497(24)	0.0418(5)	0.7067(12)	0.387	-
15	33	0.39698(23)	0.0347(8)	0.67(4)	0.398	0.40(5)
16	40	0.4076(5)	0.0284(4)	0.6501(9)	0.408	0.57(5)
17	49	0.42011(24)	0.0233(6)	0.6204(9)	0.419	0.78(5)
18	58	0.42858(28)	0.0195(4)	0.5976(9)	0.428	0.97(4)
19	68	0.43586(4)	0.01639(5)	0.58265(13)	0.435	1.204(10)
20	80	0.44415(22)	0.01405(10)	0.5625(8)	0.443	1.49(4)
21	92	0.4497(4)	0.01192(19)	0.5500(6)	0.448	1.79(5)
22	106	0.45604(25)	0.01038(12)	0.533(4)	0.454	2.13(4)
23	121	0.46150(18)	0.00907(6)	0.5192(5)	0.460	2.44(4)
24	138	0.46711(22)	0.00794(4)	0.507(8)	0.465	2.79(4)
25	156	0.47172(10)	0.00706(8)	0.4961(7)	0.470	3.14(4)
26	175	0.47568(26)	0.006283(29)	0.4864(27)	0.474	3.54(5)
27	196	0.47977(12)	0.005618(27)	0.4782(29)	0.478	3.94(4)
28	219	0.48390(19)	0.005092(30)	0.468(5)	0.482	4.31(4)
29	243	0.48727(14)	0.00461(4)	0.4595(15)	0.486	4.72(5)
30	270	0.49121(7)	-	0.4508(18)	0.490	5.13(5)
32	327	0.497118(9)	0.0035310(20)	0.43602(8)	0.496	6.020(24)
35	428	0.505299(11)	0.0027926(24)	0.4151(7)	0.504	7.471(30)
38	548	0.512368(16)	0.0022635(14)	0.3967(5)	0.511	8.907(28)
40	640	0.516700(7)	0.0019920(10)	0.3851(5)	0.516	9.94(4)
45	911	0.525666(7)	0.0014889(5)	0.3598(8)	0.525	12.67(5)
50	1250	0.533109(16)	0.0011531(9)	0.3376(6)	0.533	15.61(5)
55	1663	0.539281(7)	0.0009175(5)	0.3187(6)	0.539	18.67(6)
60	2160	0.544640(8)	0.0007462(6)	0.3017(6)	0.544	21.88(5)
65	2746	0.549220(13)	0.0006183(8)	0.299(4)	-	-
70	3430	0.553237(12)	0.00051927(24)	0.2734(11)	0.553	28.98(13)
80	5120	0.559925(11)	0.0003807(10)	0.2503(13)	0.560	36.27(10)
90	7290	0.56524(5)	0.0002892(8)	0.2318(20)	0.565	44.43(7)
100	10000	0.569727(25)	0.0002282(4)	0.2147(20)	0.570	51.90(21)

Table A.3.: Condensation of 3D Lennard-Jones gas with fixed density $\rho = 10^{-2}$. The condensation temperature T_c and rounding ΔT are obtained from the specific heat peak and the fraction $\langle\tilde{\eta}\rangle_{T_c}$ of particles in the largest cluster is determined at T_c. The equal height temperature T_{eqh} is estimated from a binary-search histogram reweighting scheme and the free-energy barrier is estimated at this temperature. Error bars, if given, are obtained from jackknife error analysis.

N	T_c	ΔT	$\langle\tilde{\eta}\rangle_{T_c}$	T_{eqh}	$\beta_{eqh}\Delta F$
4	0.24597(28)	0.0878(4)	1.0(0)	-	-
6	0.2950(4)	0.0646(4)	0.9421(4)	-	-
8	0.32670(26)	0.0523(4)	0.9040(5)	0.309	2.94(4)
10	0.34942(27)	0.04379(28)	0.9082(4)	0.340	1.91(4)
12	0.3666(4)	0.03751(25)	0.8823(4)	0.360	1.53(4)
14	0.3808(4)	0.03276(16)	0.8630(6)	0.376	1.50(4)
16	0.39282(18)	0.02827(26)	0.8485(6)	0.388	1.64(4)
18	0.40299(16)	0.02502(14)	0.8343(7)	0.399	1.80(4)
20	0.4118(5)	0.02236(14)	0.8382(6)	0.408	1.97(4)
22	0.4197(4)	0.02053(20)	0.8243(6)	0.415	2.16(4)
24	0.42657(25)	0.01881(6)	0.8136(6)	0.422	2.32(4)
26	0.43265(28)	0.01738(15)	0.8043(6)	0.428	2.49(4)
28	0.43784(21)	0.01608(8)	0.7969(7)	0.434	2.68(4)
30	0.44324(27)	0.01508(5)	0.7885(6)	0.439	2.87(5)
32	0.44772(13)	0.01419(11)	0.7801(7)	0.443	3.04(4)
36	0.45589(20)	0.01257(6)	0.7698(7)	0.452	3.39(5)
40	0.46316(12)	0.01133(6)	0.7592(7)	0.459	3.77(5)
44	0.46922(15)	0.01037(4)	0.749(6)	0.465	4.04(5)
48	0.47498(21)	0.00957(4)	0.740(13)	0.471	4.39(5)
64	0.49208(6)	0.007349(22)	0.711(14)	0.489	5.42(5)
80	0.50485(20)	0.00607(4)	0.688(4)	0.502	6.39(7)
96	0.51428(13)	0.005152(28)	0.6716(6)	0.512	7.26(5)
112	0.52200(8)	0.004532(12)	0.6539(18)	0.520	8.05(5)
128	0.52827(11)	0.004031(19)	0.641(6)	0.526	8.82(5)
160	0.53858(14)	0.003359(9)	0.617(4)	0.537	10.16(6)
192	0.54635(13)	0.002881(12)	0.5999(23)	0.545	11.44(5)
224	0.55264(9)	0.002549(8)	0.583(4)	0.551	12.54(5)
256	0.55784(13)	0.002282(7)	0.570(4)	0.556	13.66(5)
320	0.56620(10)	0.001913(9)	0.546(4)	0.565	15.50(8)
384	0.57252(5)	0.001653(5)	0.5273(28)	0.571	17.24(6)
448	0.57767(19)	0.001462(8)	0.512(4)	0.577	18.85(7)
512	0.58199(9)	0.001317(4)	0.499(4)	0.581	20.25(8)

Table A.4.: Aggregation of 3D flexible lattice polymers at fixed density $\rho = 10^{-2}$. The aggregation temperature T_{agg} and rounding ΔT are obtained from the specific heat peak and the fraction $\langle\tilde{\eta}\rangle_{T_{\mathrm{agg}}}$ of monomers in the largest aggregate is determined at T_{agg}. The equal height temperature T_{eqh} is estimated from a binary-search histogram reweighting scheme and the free-energy barrier is estimated at this temperature. Error bars, if given, are obtained from jackknife error analysis.

L	N	T_{c}	ΔT_{c}	$\langle\tilde{\eta}\rangle_{T_{\mathrm{c}}}$	T_{eqh}	$\beta_{\mathrm{eqh}}\Delta F$
14	2	0.9026(18)	0.9246(24)	1.0(0)	-	-
16	3	1.0235(21)	0.610(15)	1.0(0)	-	-
18	4	1.0890(18)	0.4241(30)	0.90994(27)	-	-
20	6	1.1912(7)	0.3223(20)	0.92507(19)	-	-
22	8	1.2525(5)	0.2597(28)	0.87910(25)	-	-
24	11	1.3223(6)	0.2096(17)	0.85161(21)	-	-
26	14	1.367(5)	0.1709(22)	0.8006(10)	1.382	-
28	17	1.39654(21)	0.1419(10)	0.78641(20)	1.407	-
32	25	1.4540(10)	0.0961(30)	0.71801(29)	1.458	-
32	25	1.4542(7)	0.0961(30)	0.71801(29)	1.458	-
36	36	1.5055(4)	0.0669(6)	0.67321(26)	1.506	0.52(4)
40	49	1.5406(6)	0.04706(25)	0.63199(27)	1.539	0.86(4)
44	66	1.57440(28)	0.0344(5)	0.59214(28)	1.572	1.29(5)
48	85	1.59730(20)	0.02587(17)	0.55979(24)	1.594	1.73(5)
52	108	1.61796(18)	0.01994(10)	0.5323(30)	1.615	2.29(6)
56	135	1.6358(7)	0.0159(4)	0.5099(10)	1.633	2.89(10)
60	166	1.6515(6)	0.01285(17)	0.488(7)	1.648	3.42(10)
62	183	1.6578(8)	0.01151(17)	0.483(4)	1.655	3.81(10)
64	202	1.6652(5)	0.01044(18)	0.474(6)	1.662	4.03(8)
68	242	1.6768(6)	0.00879(19)	0.457(7)	1.674	4.78(11)
72	287	1.6870(4)	0.00745(7)	0.443(6)	1.685	5.48(10)
76	338	1.69667(26)	0.00641(7)	0.429(4)	1.694	6.21(10)
80	394	1.7048(5)	0.00551(7)	0.420(5)	1.703	7.24(16)
84	456	1.71270(19)	0.00488(4)	0.406(4)	1.711	7.77(11)
88	524	1.7193(4)	0.00428(6)	0.397(6)	1.718	8.87(27)

Table A.5.: Aggregation of 3D flexible bead-spring polymers at fixed density $\rho = 10^{-2}$. The aggregation temperature T_{agg} and rounding ΔT are obtained from the specific heat peak and the fraction $\langle\tilde{\eta}\rangle_{T_{\mathrm{agg}}}$ of monomers in the largest aggregate is determined at T_{agg}. The equal height temperature T_{eqh} is estimated from a binary-search histogram reweighting scheme and the free-energy barrier is estimated at this temperature. Error bars, if given, are obtained from jackknife error analysis.

N	T_{c}	ΔT_{c}	$\langle\tilde{\eta}\rangle_{T_{\mathrm{c}}}$	T_{eqh}	$\beta_{\mathrm{eqh}}\Delta F$
4	1.1004(20)	0.2175(29)	0.9131(13)	1.112	0.19(4)
6	1.1683(22)	0.1345(27)	0.9305(14)	1.175	0.24(6)
8	1.2127(17)	0.0993(24)	0.892(4)	1.216	0.25(5)
10	1.243(4)	0.0786(26)	0.890(5)	1.244	0.39(5)
12	1.2661(16)	0.0627(15)	0.834(4)	1.266	0.56(6)
14	1.2829(28)	0.0529(26)	0.8157(20)	1.282	0.73(6)
16	1.2980(22)	0.0461(15)	0.8008(21)	1.297	0.83(5)
18	1.3111(17)	0.0407(19)	0.7837(18)	1.309	1.02(7)
20	1.3217(19)	0.036(4)	0.7755(22)	1.319	1.17(7)
24	1.3383(18)	0.0291(8)	0.7567(19)	1.336	1.45(7)
28	1.351(4)	0.0244(12)	0.7414(19)	1.348	1.75(7)
32	1.3637(20)	0.0206(13)	0.7350(26)	1.361	2.06(8)
36	1.3717(16)	0.0182(5)	0.7191(20)	1.369	2.37(10)
40	1.3804(30)	0.0165(6)	0.71(4)	1.377	2.56(9)
48	1.3955(11)	0.01366(22)	0.6831(25)	1.393	3.05(11)
56	1.4037(11)	0.0115(7)	0.6733(17)	1.401	3.58(9)
60	1.4086(4)	0.01078(16)	0.6672(19)	1.406	3.69(8)
64	1.4109(12)	0.01005(27)	0.660(21)	-	-

Bibliography

[1] M. Barber, in *Phase Transitions and Critical Phenomena*, Vol. 8, edited by C. Domb and J. Lebowitz (Academic, NY, 1983).

[2] P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, 1953).

[3] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, 1979).

[4] F. Chiti and C. M. Dobson, Annu. Rev. Biochem. **75**, 333 (2006).

[5] F. Müller-Plathe, ChemPhysChem **3**, 754 (2002).

[6] W. Janke, in *Order, Disorder and Criticality: Advanced Problems of Phase Transition Theory*, Vol. 3, edited by Y. Holovatch (World Scientific, Singapore, 2012) pp. 93–166.

[7] W. Janke (Editor), *Rugged free energy landscapes: Common computational approaches to spin glasses, structural glasses and biological macromolecules*, Lect. Notes. Phys., Vol. 736 (Springer, Berlin, 2008).

[8] M. E. Fisher, Rep. Prog. Phys. **30**, 615 (1967).

[9] K. Binder and D. Stauffer, Adv. Phys. **25**, 343 (1976).

[10] M. Biskup, L. Chayes, and R. Kotecký, Europhys. Lett. **60**, 32 (2002).

[11] M. Biskup, L. Chayes, and R. Kotecký, J. Stat. Phys. **116**, 175 (2003).

[12] T. Neuhaus and J. Hager, J. Stat. Phys. **113**, 47 (2003).

[13] K. Binder, Physica A **319**, 99 (2003).

[14] K. Binder and M. H. Kalos, J. Stat. Phys. **22**, 363 (1980).

[15] A. Nußbaumer, E. Bittner, T. Neuhaus, and W. Janke, Europhys. Lett. **75**, 716 (2006).

[16] A. Nußbaumer, E. Bittner, and W. Janke, Phys. Rev. E **77**, 041109 (2008).

[17] J. Zierenberg, M. Wiedenmann, and W. Janke, J. Phys: Conf. Ser. **510**, 012017 (2014).

[18] L. G. MacDowell, P. Virnau, M. Müller, and K. Binder, J. Chem. Phys. **120**, 5293 (2004).

[19] M. Schrader, P. Virnau, and K. Binder, Phys. Rev. E **79**, 061104 (2009).

[20] S. Martinos, A. Malakis, and I. Hadjiagapiou, Physica A **384**, 368 (2007).

[21] J. Zierenberg and W. Janke, to appear in Phys. Rev. E (2015), in print.

[22] T. D. Lee and C. N. Yang, Phys. Rev. **87**, 410 (1952).

[23] L. Onsager, Nuovo Cimento, Suppl. **6**, 261 (1949).

[24] C. N. Yang, Phys. Rev. **85**, 808 (1952).

[25] W. P. Orrick, B. G. Nickel, A. J. Guttmann, and J. H. H. Perk, Phys. Rev. Lett. **86**, 4120 (2001).

[26] W. Orrick, B. Nickel, A. Guttmann, and J. Perk, J. Stat. Phys. **102**, 795 (2001).

[27] B. Nickel, J. Phys. A: Math. Gen. **32**, 3889 (1999).

[28] B. Nickel, J. Phys. A: Math. Gen. **33**, 1693 (2000).

[29] S. Boukraa, A. J. Guttmann, S. Hassani, I. Jensen, J. M. Maillard, B. Nickel, and N. Zenine, J. Phys. A: Math. Theor. **41**, 455202 (2008).

[30] Y. Chan, A. J. Guttmann, B. G. Nickel, and J. H. H. Perk, J. Stat. Phys. **145**, 549 (2011).

[31] K. Leung and R. K. P. Zia, J. Phys. A: Math. Gen. **23**, 4593 (1990).

[32] J. Adler, Phys. Rev. B **36**, 2473 (1987).

[33] M. Hasenbusch, S. Meyer, and M. Pütz, J. Stat. Phys. **85**, 383 (1996).

[34] M. F. Sykes, J. W. Essam, and D. S. Gaunt, J. Math. Phys. **6**, 283 (1965).

[35] G. Bhanot, M. Creutz, and J. Lacki, Phys. Rev. Lett. **69**, 1841 (1992).

[36] L. J. Shaw and M. E. Fisher, Phys. Rev. A **39**, 2189 (1989).

[37] H. Arisue, Phys. Lett. B **313**, 187 (1993).

[38] M. Fisher and J.-H. Chen, J. Phys. **46**, 1645 (1985).

[39] Y. Imry, Phys. Rev. B **21**, 2042 (1980).

[40] E. Bittner, A. Nußbaumer, and W. Janke, Nucl. Phys. B **820**, 694 (2009).

[41] F. Schmitz, P. Virnau, and K. Binder, Phys. Rev. E **87**, 053302 (2013).

[42] J. Zierenberg, M. Mueller, P. Schierz, M. Marenz, and W. Janke, J. Chem. Phys. **141**, 114908 (2014).

[43] A. Trokhymchuk and J. Alejandre, J. Chem. Phys. **111**, 8510 (1999).

[44] G. A. Chapela, G. Saville, S. M. Thompson, and J. S. Rowlinson, J. Chem. Soc., Faraday Trans. 2 **73**, 1133 (1977).

[45] V. Privman and J. Rudnick, J. Stat. Phys. **60**, 551 (1990).

[46] C. Borgs and R. Kotecký, J. Stat. Phys. **79**, 43 (1995).

[47] C. Borgs, R. Kotecký, and I. Medved, J. Stat. Phys. **109**, 67 (2002).

[48] A. Nußbaumer, E. Bittner, and W. Janke, Prog. Theor. Phys. Suppl. **184**, 400 (2010).

[49] D. H. E. Gross, *Microcanonical Thermodynamics: Phase Transitions in "Small" Systems* (World Scientific, Singapore, 2001).

[50] W. Janke, Nucl. Phys. B (Proc. Suppl.) **63**, 631 (1998).

[51] C. Junghans, M. Bachmann, and W. Janke, Phys. Rev. Lett. **97**, 218103 (2006).

[52] C. Junghans, M. Bachmann, and W. Janke, J. Chem. Phys. **128**, 085103 (2008).

[53] C. Junghans, M. Bachmann, and W. Janke, Europhys. Lett. **87**, 40002 (2009).

[54] M. Möddel, W. Janke, and M. Bachmann, Phys. Chem. Chem. Phys. **12**, 11548 (2010).

[55] S. Schnabel, D. T. Seaton, D. P. Landau, and M. Bachmann, Phys. Rev. E **84**, 011127 (2011).

[56] J. R. Ray, Phys. Rev. A **44**, 4061 (1991).

[57] R. Lustig, J. Chem. Phys. **109**, 8816 (1998).

[58] V. Martin-Mayor, Phys. Rev. Lett. **98**, 137207 (2007).

[59] P. Schierz, J. Zierenberg, and W. Janke, Leipzig preprint (2015), submitted.

[60] R. Ni, S. Abeln, M. Schor, M. A. C. Stuart, and P. G. Bolhuis, Phys. Rev. Lett. **111**, 058101 (2013).

[61] A. Rahman, Phys. Rev. **136**, A405 (1964).

[62] C. Tegeler, R. Span, and W. Wagner, J. Phys. Chem. Ref. Data **28**, 779 (1999).

[63] W. Paul, T. Strauch, F. Rampf, and K. Binder, Phys. Rev. E **75**, 060801 (2007).

[64] C. Junghans, W. Janke, and M. Bachmann, Comp. Phys. Comm. **182**, 1937 (2011).

[65] M. Mueller, J. Zierenberg, M. Marenz, P. Schierz, and W. Janke, to appear in Physics Procedia (2015), in print.

[66] T. Vogel, M. Bachmann, and W. Janke, Phys. Rev. E **76**, 061803 (2007).

[67] N. Madras and A. D. Sokal, J. Stat. Phys. **50**, 109 (1988).

[68] A. Milchev, A. Bhattacharya, and K. Binder, Macromolecules **34**, 1881 (2001).

[69] S. Schnabel, M. Bachmann, and W. Janke, J. Chem. Phys. **131**, 124904 (2009).

[70] S. Schnabel, W. Janke, and M. Bachmann, J. Comp. Phys. **230**, 4454 (2011).

[71] W. Janke, in *Computational Many-Particle Physics*, Lect. Notes. Phys., Vol. 739, edited by H. Fehske, R. Schneider, and A. Weiße (Springer, Berlin, 2008) pp. 79–140.

[72] M. Möddel, W. Janke, and M. Bachmann, Macromolecules **44**, 9013 (2011).

[73] S. Karalus, W. Janke, and M. Bachmann, Phys. Rev. E **84**, 031803 (2011).

[74] M. Möddel, *Statistical Equilibrium Behaviour of Finite Polymers Near Attractive Substrates*, Ph.D. thesis, Universität Leipzig (2012).

[75] M. Möddel, W. Janke, and M. Bachmann, Phys. Rev. Lett. **112**, 148303 (2014).

[76] J. Kierfeld and R. Lipowsky, Europhys. Lett. **62**, 285 (2003).

[77] B. Schott, Master's thesis, Universität Leipzig (in preparation).

[78] D. F. Parsons and D. R. Williams, Phys. Rev. E **74**, 041804 (2006).

[79] D. F. Parsons and D. R. Williams, J. Chem. Phys. **124**, 221103 (2006).

[80] D. Seaton, T. Wüst, and D. Landau, Phys. Rev. E **81**, 011802 (2010).

[81] J. Gross, T. Neuhaus, T. Vogel, and M. Bachmann, J. Chem. Phys. **138**, 074905 (2013).

[82] P. Young, arXiv:1210.3781 (2012).

[83] S. Doniach, T. Garel, and H. Orland, J. Chem. Phys. **105**, 1601 (1996).

[84] U. Bastolla and P. Grassberger, J. Stat. Phys. **89**, 1061 (1997).

[85] A. Maritan, C. Micheletti, A. Trovato, and J. R. Banavar, Nature **406**, 287 (2000).

[86] J. R. Banavar and A. Maritan, Rev. Mod. Phys. **75**, 23 (2003).

[87] T. Vogel, T. Neuhaus, M. Bachmann, and W. Janke, Europhys. Lett. **85**, 10003 (2009).

[88] T. Vogel, T. Neuhaus, M. Bachmann, and W. Janke, Phys. Rev. E **80**, 011802 (2009).

[89] T. Vogel, T. Neuhaus, M. Bachmann, and W. Janke, Eur. Phys. J. E **30**, 7 (2009).

[90] D. T. Seaton, S. Schnabel, D. P. Landau, and M. Bachmann, Phys. Rev. Lett. **110**, 028103 (2013).

[91] M. Marenz and W. Janke, Physics Procedia **57**, 53 (2014).

[92] S. Auer, C. M. Dobson, M. Vendruscolo, and A. Maritan, Phys. Rev. Lett. **101**, 258101 (2008).

[93] S. Auer and D. Kashchiev, Phys. Rev. Lett. **104**, 168105 (2010).

[94] J. R. Banavar, T. X. Hoang, J. H. Maddocks, A. Maritan, C. Poletto, A. Stasiak, and A. Trovato, Proc. Natl. Acad. Sci. U.S.A. **104**, 17283 (2007).

[95] J. Zierenberg and W. Janke, Europhys. Lett. **109**, 28002 (2015).

[96] J. Kierfeld, T. Kühne, and R. Lipowsky, Phys. Rev. Lett. **95**, 038102 (2005).

[97] C. Heussinger, F. Schüller, and E. Frey, Phys. Rev. E **81**, 021904 (2010).

[98] G. M. Grason and R. F. Bruinsma, Phys. Rev. Lett. **99**, 098101 (2007).

[99] M. Turner, R. Briehl, F. Ferrone, and R. Josephs, Phys. Rev. Lett. **90**, 128103 (2003).

[100] J. T. Giurleo, X. He, and D. S. Talaga, J. Mol. Biol. **381**, 1332 (2008), `http://talaga.rutgers.edu/research/images/Amyloid%20Mechanisms-2.jpg`.

[101] T. Vogel and M. Bachmann, Phys. Rev. Lett. **104**, 198302 (2010).

[102] Y. Yoshimura, Y. Lin, H. Yagi, Y.-H. Lee, H. Kitayama, K. Sakurai, M. So, H. Ogi, H. Naiki, and Y. Goto, Proc. Natl. Acad. Sci. U.S.A. **109**, 14446 (2012).

[103] O. Kratky and G. Porod, Rec. Trav. Chim. Pays-Bas **68**, 1106 (1949).

[104] O. Kratky and G. Porod, J. Colloid Sci. **4**, 35 (1949).

[105] S. Schöbl, *Macromolecules in Disordered Environments: From Flexible to Semiflexible Polymers*, Ph.D. thesis, Universität Leipzig (2013).

[106] M. P. Allen, G. T. Evans, D. Frenkel, and B. Mulder, Adv. Chem. Phys. **86**, 1 (1993).

[107] T. Iwataki, S. Kidoaki, T. Sakaue, K. Yoshikawa, and S. S. Abramchuk, J. Chem. Phys. **120**, 4004 (2004).

[108] M. Claessens, C. Semmrich, L. Ramos, and A. Bausch, Proc. Natl. Acad. Sci. U.S.A. **105**, 8819 (2008).

[109] L. Haviv, N. Gov, Y. Ideses, and A. Bernheim-Groswasser, Eur. Biophys. J. **37**, 447 (2008).

[110] L. C. Serpell, Biochimica et Biophysica Acta **1502**, 16 (2000).

[111] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[112] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **57**, 2607 (1986).

[113] C. J. Geyer, in *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, edited by E. M. Keramidas (Interface Foundation, Fairfax Station, VA, 1991) pp. 156–163.

[114] K. Hukushima and K. Nemoto, J. Phys. Soc. Japan **65**, 1604 (1996).

[115] U. H. E. Hansmann, Chem. Phys. Lett. **281**, 140 (1997).

[116] D. J. Earl and M. W. Deem, Phys. Chem. Chem. Phys. **7**, 3910 (2005).

[117] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).

[118] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).

[119] W. Janke, Physica A **254**, 164 (1998).

[120] J. Zierenberg, M. Marenz, and W. Janke, Comp. Phys. Comm. **184**, 1155 (2013).

[121] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, 2000).

[122] B. A. Berg, *Markov Chain Monte Carlo Simulations and Their Statistical Analysis* (World Scientific Publishing, 2004).

[123] R. H. Swendsen, *An Introduction to Statistical Mechanics and Thermodynamics* (Oxford University Press, 2012).

[124] M. Marenz, J. Zierenberg, H. Arkın, and W. Janke, Condens. Matter Phys. **15**, 43008 (2012).

[125] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).

[126] F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).

[127] "Bézier curve," in *Encyclopedia of Mathematics.*, `http://www.encyclopediaofmath.org/index.php?title=B%C3%A9zier_curve&oldid=23190`.

[128] J. Zierenberg, M. Marenz, and W. Janke, Physics Procedia **53**, 55 (2014).

[129] W. Janke, B. A. Berg, and M. Katoot, Nucl. Phys. B **382**, 649 (1992).

[130] A. Nußbaumer, E. Bittner, T. Neuhaus, and W. Janke, Physics Procedia **7**, 52 (2010).

[131] M. Bachmann, H. Arkın, and W. Janke, Phys. Rev. E **71**, 031906 (2005).

[132] M. Lal, Molec. Phys. **17**, 57 (1969).

[133] B. MacDonald, N. Jan, D. L. Hunter, and M. O. Steinitz, J. Phys. A: Math. Gen. **18**, 2627 (1985).

[134] N. C. Karayiannis, V. G. Mavrantzas, and D. N. Theodorou, Phys. Rev. Lett. **88**, 105503 (2002).

[135] J. Baschnagel, J. P. Wittmer, and H. Meyer, in *Computational Soft Matter: From Synthetic Polymers to Proteins*, NIC Series, Vol. 23, edited by N. Attig, K. Binder, H. Grubmller, and K. Kremer (2004) pp. 83–140.

[136] K. Binder and W. Paul, Macromolecules **41**, 4537 (2008).

[137] R. Auhl, R. Everaers, G. S. Grest, K. Kremer, and S. J. Plimpton, J. Chem. Phys. **119**, 12718 (2003).

[138] D. Reith and P. Virnau, Comp. Phys. Comm. **181**, 800 (2010).

[139] B. Efron, *The Jackknife, the Bootstrap and Other Resampling Plans* (Society for Industrial and Applied Mathematics, Philadelphia, 1982).