

## Collapse transition of a Lennard-Jones polymer

Stefan Schnabel  and Wolfhard Janke 

*Institut für Theoretische Physik, Universität Leipzig, IPF 231101, 04081 Leipzig, Germany*

 (Received 5 September 2023; accepted 16 November 2023; published 18 December 2023)

Using the recently introduced parsimonious Metropolis Monte Carlo algorithm, bead-stick polymers both with infinite-range Lennard-Jones interaction and with truncation are simulated. The focus lies on determining the Boyle temperature for long chains with thousands of repeat units and on testing for theoretically predicted logarithmic corrections. Subsequently the behavior at the infinite-chain transition temperature, i.e., the  $\Theta$  temperature, is studied for chains with up to  $N = 32\,768$  repeat units by investigation of the scaling of the end-to-end distance, the radius of gyration, the specific heat, and their derivatives with  $N$ .

DOI: [10.1103/PhysRevE.108.064502](https://doi.org/10.1103/PhysRevE.108.064502)

### I. INTRODUCTION

Flexible polymers in dilute solution undergo a coil-globule or  $\Theta$  transition when the solvent is changed from good to bad or the temperature changes from high to low [1,2]. The critical temperature is the infinite-chain limit of the Boyle temperature, the temperature at which the second virial coefficient vanishes [3]. This transition is expected to be of second order with the “size” of the polymer in the form of the end-to-end distance changing continuously. At the critical point the polymer behaves on large enough scales like an ideal chain or Gaussian random walk. At higher and lower temperatures the chain resembles a self-avoiding walk and a collapsed globule, respectively. Being a tricritical point [4], the upper critical dimension of the  $\Theta$  transition is  $d_c = 3$  [1,2] and logarithmic corrections to scaling should be present. They have been calculated to first order [5–7] and later to second order [8] by means of renormalization-group methods.

However, these predicted logarithmic corrections have not been verified by means of Monte Carlo (MC) simulations, which is due to the fact that the logarithmic asymptotic scaling manifests clearly only for very large systems, which are difficult to handle. In the past, chains with thousands of repeat units could only be investigated in the form of self-attracting lattices walks [9,10], and off-lattice systems with only some dozens [11] or few hundreds [12–14] of monomers have been realized. Due to these limitations, studies determining  $\Theta$  temperatures are few and far between even though there is great interest in transitions of single-polymer chains and their phase diagram. For more recent works see, for instance, [15,16]. In an attempt to address the problem we recently introduced a MC method [17] that substantially accelerates the simulation of long polymer chains with continuous degrees of freedom and extends the length of chains that can be treated by at least an order of magnitude. In this study we now employ this method to investigate the  $\Theta$  transition of a Lennard-Jones polymer.

The paper is organized as follows: Sec. II is devoted to a review of the model and geometric observables. In Sec. III we discuss the MC methods we use, how the second virial coefficient is measured, and how averages of observables based

on energy can be measured very efficiently. The outcome of our simulations is then presented and discussed in Sec. IV. In Sec. V we finish with a summary and some concluding remarks.

### II. MODEL

We consider a bead-stick chain  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$  in three dimensions where the monomers  $\mathbf{x}_i$  interact pairwise via a 12-6 Lennard-Jones (LJ) interaction:

$$U_{\text{LJ}}(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right]. \quad (1)$$

The chain is fully flexible, but the bonds connecting adjacent monomers are of fixed length:

$$|\mathbf{x}_k - \mathbf{x}_{k-1}| = b. \quad (2)$$

The monomer diameter is adjusted to  $\sigma = 2^{-1/6}b$  such that  $U_{\text{LJ}}(b) = -\epsilon$ .

As a variation of the model we will also use the truncated and shifted LJ potential

$$U_c(r) = U_{\text{LJ}}(\min(r, r_c)) - U_{\text{LJ}}(r_c), \quad (3)$$

where  $r_c$  is the cutoff distance. The shapes of the resulting potentials are depicted in Fig. 1. The total energy function for an isolated chain reads

$$E = \sum_{i=1}^{N-1} \sum_{j=i+1}^N U(|\mathbf{x}_i - \mathbf{x}_j|). \quad (4)$$

Two geometric observables are commonly considered: the end-to-end distance  $r_{\text{ee}} = |\mathbf{x}_N - \mathbf{x}_1|$  and the radius of gyration

$$r_{\text{gyr}}^2 = \frac{1}{N} \sum_{i=1}^N \left( \mathbf{x}_i - \frac{1}{N} \sum_{j=1}^N \mathbf{x}_j \right)^2. \quad (5)$$

In the following we will present distances in units of  $b$  and temperatures in units of  $k_B/\epsilon$ .

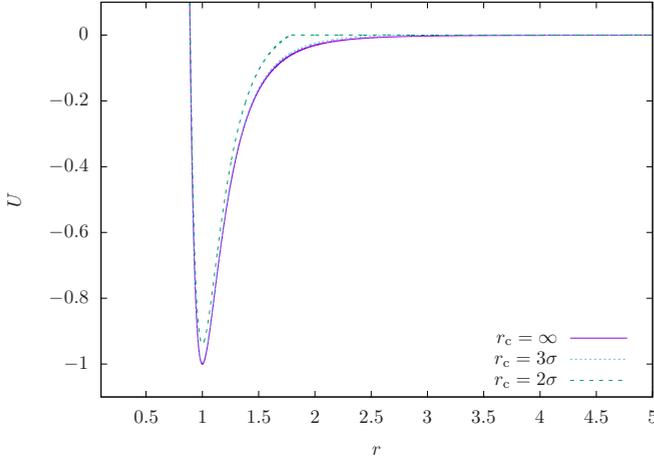


FIG. 1. LJ potentials (1) and (3) without truncation and with truncation at  $r = r_c$ .

### III. METHODS

#### A. Conformational updates

To generate new conformations we apply three different MC moves. First, the pivot move [18] is used. It randomly picks a monomer as a fulcrum and rotates all monomers on the smaller side,

$$\mathbf{x}'_i = \mathcal{R}(\mathbf{x}_i - \mathbf{x}_k) + \mathbf{x}_k, \quad (6)$$

for all  $i > k$  if  $k > N/2$  or for all  $i < k$  if  $k \leq N/2$ , respectively. Here  $\mathcal{R}$  is a random rotation matrix. Second, we use a bond rotation, where a single bond  $\mathbf{b}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$  is randomly assigned a new direction while all other bonds remain the same. Consequently

$$\mathbf{x}'_i = \mathbf{x}_i + \mathbf{b}'_i - \mathbf{b}_k \quad \text{for all } i > k \quad (7)$$

if  $k > N/2$  or

$$\mathbf{x}'_i = \mathbf{x}_i - \mathbf{b}'_i + \mathbf{b}_k \quad \text{for all } i < k \quad (8)$$

if  $k \leq N/2$ , respectively.

As we will discuss below, in order to determine the second virial coefficient, it is necessary to simulate two interacting chains. In this situation we additionally employ a generalized crank-shaft move: We select two monomers at positions  $\mathbf{x}_k$  and  $\mathbf{x}_l$ , determine the axis through both of them, and rotate all monomers in between:

$$\mathbf{x}'_i = \mathcal{R}(\mathbf{x}_i - \mathbf{x}_k) + \mathbf{x}_k \quad \text{for all } k < i < l, \quad (9)$$

where  $\mathcal{R}$  is a random rotation with the constraint  $\mathbf{x}_k - \mathbf{x}_l = \mathcal{R}(\mathbf{x}_k - \mathbf{x}_l)$ . We bias the selection of monomers such that predominantly small parts of the polymer are modified. To this end two random numbers  $\xi_1, \xi_2 \in [0, 1)$  are used to first assign the length of the rotating segment  $w = 2 + \lfloor \xi_1^2 N/3 \rfloor \geq 2$ . Here  $\lfloor \cdot \rfloor$  is the floor function and the choice of  $\xi_1^2$  instead of just  $\xi_1$  increases the chance of selecting smaller values for  $w$ . Then we find the limiting left and right monomer indices by setting  $k = 1 + \lfloor \xi_2(N - w) \rfloor$  leading to  $k \in [1, N - w]$  and  $l = k + w$ .

Since this study focuses on the collapse transition, i.e., on the two phases at medium and high temperature, the more

elaborate moves designed to sample the crystal-like low-temperature state [19] are not needed.

#### B. Binary tree data structure

Clisby [20] introduced a binary tree data structure for the efficient sampling of self-avoiding walks on lattices which allowed the determination of scaling exponents with very high accuracy [20–22]. Recently we generalized the method for the simulation of off-lattice polymers [17]. In the data tree that it employs, each node relates to a piece of the chain. The root node represents the entire polymer, its children are the polymer's two halves, each child node of the root's children stands for a quarter, and so on. With each level down the tree, the number of nodes doubles while the amount of monomers in each node is reduced by a factor of two. The nodes do not store the actual monomer positions but rather the information of a simple geometrical body that contains all the node's monomers. While on a cubic lattice a natural choice for this body is a cuboid, in the present case of continuous degrees of freedom we use spheres. For the tree's leaves, i.e., for individual monomers, these spheres are situated at the monomer position  $\mathbf{x}_k$ , and they have radius zero. At higher levels, spheres are constructed recursively by finding the smallest sphere that contains the spheres of the child nodes.

If a node  $A$  contains  $n_A$  monomers within a sphere of radius  $r_A$  around the center point  $\mathbf{c}_A$ , then the interaction  $E_{AB}$  between two nodes  $A$  and  $B$  with disjoint sets of monomers can be estimated. If the two spheres are distant, this can be done with good precision and little computational effort, and as we will discuss further below an efficient MC method can be based on this idea. Defining the extremal distances

$$d_{AB}^{\min} = |\mathbf{c}_A - \mathbf{c}_B| - r_A - r_B, \quad (10)$$

$$d_{AB}^{\max} = |\mathbf{c}_A - \mathbf{c}_B| + r_A + r_B, \quad (11)$$

we find that if  $d_{AB}^{\min} \geq 2^{1/6}\sigma$ ,

$$E_{AB} \in [n_A n_B U(d_{AB}^{\min}), n_A n_B U(d_{AB}^{\max})], \quad (12)$$

since under this condition the potentials (1) and (3) are monotonically increasing throughout  $[d_{AB}^{\min}, d_{AB}^{\max}]$ . If the uncertainty is too large, i.e., if the interval of possible values for  $E_{AB}$  is too wide and narrower boundaries are desired, it is possible to “increase the resolution” by descending one branch of the tree. If we label the child nodes of  $A$  as  $A_1$  and  $A_r$  we obtain

$$E_{AB} \in [n_{A_1} n_B U(d_{A_1 B}^{\min}) + n_{A_r} n_B U(d_{A_r B}^{\min}), n_{A_1} n_B U(d_{A_1 B}^{\max}) + n_{A_r} n_B U(d_{A_r B}^{\max})] \quad (13)$$

if  $d_{A_1 B}^{\min} \geq 2^{1/6}\sigma$  and  $d_{A_r B}^{\min} \geq 2^{1/6}\sigma$ . Of course, instead of  $A$ , node  $B$  can be replaced by its children as well. Intuitive heuristics are to split the node that contains the greater number of monomers or the node with the greater sphere, both of which are sound strategies leading to similar performance.

The internal energy of any node  $A$  is recursively given by

$$E_A = E_{A_1 A_r} + E_{A_1} + E_{A_r}, \quad (14)$$

and the total energy (4) is the internal energy of the tree's root node. In virtually all cases the straightforward recursive evaluation of Eq. (14) will not directly lead to an estimate of

$E_A$  since many adjacent nodes will have overlapping spheres or minimal distances smaller than  $2^{1/6}\sigma$ . Hence, applications of the refinement described in Eq. (13) are required. However, even when all involved node-node interactions can be estimated initially, the uncertainty of the resulting sum will often be too large such that further improvement is necessary. We apply two techniques to control the precision.

First, one can choose to accept only estimations of node-node interactions that do not exceed a given uncertainty  $\kappa$  per monomer pair, i.e., both  $d_{CD}^{\min} \geq 2^{1/6}\sigma$  and  $U(d_{CD}^{\max}) - U(d_{CD}^{\min}) \leq \kappa$  must be true or the interaction between nodes  $C$  and  $D$  is further decomposed by splitting either  $C$  or  $D$ . This has the advantage that decisions can be made locally without reference to the global picture such that a recursive evaluation of Eq. (14) with minimal memory requirement is possible. However, if done in this way it is not possible to easily improve the precision of the sum once the calculation is completed. The entire process has to be run again with a smaller  $\kappa$ .

Alternatively, one can keep track of all node-node interactions that currently contribute to the sum and sort them according to their respective uncertainties. Thereby, it is not necessary to establish strict order; instead it suffices to group the interactions such that the uncertainties of interactions in each group are roughly similar, e.g., do not differ by more than a factor of two. It is then possible to repeatedly select an interaction  $E_{CD}$  with a large uncertainty and replace it by the two node-node interactions obtained when either  $C$  or  $D$  is split, thus improving the global estimate until a satisfactory precision is reached.

The binary tree also proves very useful in the case of the truncated potential. Then the interaction between two nodes  $A$  and  $B$  can be recursively evaluated exactly:

$$E_{AB} = E_{A_1B} + E_{A_2B} \quad \text{or} \quad (15)$$

$$E_{AB} = E_{AB_1} + E_{AB_2}, \quad (16)$$

with the recursion ending if either  $d_{CD}^{\min} \geq r_c$  meaning that there is no interaction between the two nodes  $C$  and  $D$ , or if  $n_C = n_D = 1$  in which case the monomer-monomer potential is directly calculated. This procedure can be seen as an efficient way to determine which of the  $n_A n_B$  monomer pairs are closer than the cutoff radius  $r_c$  and contribute to  $E_{AB}$ .

### C. The parsimonious Metropolis algorithm

When the canonical ensemble is sampled using the Metropolis algorithm [23] the acceptance of a proposed update from state  $\mu$  to state  $\nu$  is given by

$$P_{\mu \rightarrow \nu}^{\text{acc}} = \min(1, P_\nu/P_\mu) = \min(1, e^{-\beta \Delta E}), \quad (17)$$

where  $\Delta E = E_\nu - E_\mu$  is the difference in energy between the two states and  $\beta = (k_B T)^{-1} \geq 0$  the inverse temperature. The standard method to decide whether an update is to be accepted usually involves first checking if  $\Delta E \leq 0$ , which directly implies acceptance. If on the other hand  $\Delta E > 0$ , one finds  $P_{\mu \rightarrow \nu}^{\text{acc}} \in (0, 1)$  and a decision is made by drawing a random number  $\xi \in [0, 1)$  and accepting the update if  $\xi < P_{\mu \rightarrow \nu}^{\text{acc}}$ . This procedure is especially suitable if the change in energy can be calculated very fast, while drawing a random number is

in comparison computationally expensive. A typical example for such a combination are spin flips in the (short-range) Ising model and the widely used Mersenne Twister random number generator. If on the other hand the calculation of  $\Delta E$  is the more demanding part such as in the case of long-range interacting systems, it can be beneficial to use an alternative technique: Draw the random number  $\xi \in [0, 1)$  first and determine  $\eta = -\beta^{-1} \ln \xi$ . Since an update is accepted if  $\Delta E < \eta$  and rejected otherwise, the task at hand is no longer to calculate the change in energy with maximal precision, but rather to establish whether this inequality is true or false. As we have demonstrated above, strict bounds for the energy can be established and refined very efficiently using the binary tree data structure. Further details of the algorithm can be found in [17,24].

### D. Measuring energetic observables

As we have indicated, the technique we are applying for the simulation of LJ polymers with untruncated potential does not require the exact calculation of the changes in energy associated with each MC step, and it is, therefore, very fast. However, we are, of course, interested in quantities like the average energy  $\langle E \rangle$ , the second moment  $\langle E^2 \rangle$ , and averages of products with other observables, e.g.,  $\langle r_{cc} E \rangle$  or more general averages of observables of the form  $\langle O(E, \dots) \rangle$ . In order to minimize the statistical error it is desirable to measure these observables as often as possible, yet the huge computational effort ( $\propto N^2$ ) required to calculate  $E = \sum_{i=1}^{N-1} \sum_{j=i+1}^N U(|\mathbf{x}_i - \mathbf{x}_j|)$  does not allow this. However, we can estimate  $E$  very efficiently as long as the precision is low; i.e., with much less effort we can measure an energy-like observable  $\mathcal{E}_\delta$  that fulfills the condition  $|\mathcal{E}_\delta - E| < \delta$ .

If  $\delta$  can be chosen freely, we can express the average energy for instance in the following way:

$$\langle E \rangle = \langle \mathcal{E}_{10^{-1}} \rangle + \langle \mathcal{E}_{10^{-4}} - \mathcal{E}_{10^{-1}} \rangle + \langle E - \mathcal{E}_{10^{-4}} \rangle. \quad (18)$$

The first term on the right can be seen as an imprecise estimate, the second is its correction, and the third a correction of this correction. While measuring these averages our goal should be to end up with comparable statistical uncertainties for all of them. While  $\mathcal{E}_{10^{-1}}$  has a wide distribution the many measurements that are needed for a precise determination of  $\langle \mathcal{E}_{10^{-1}} \rangle$  do not slow the simulation since the large  $\delta = 10^{-1}$  allows for a fast evaluation. Next, the distribution of  $\mathcal{E}_{10^{-4}} - \mathcal{E}_{10^{-1}} \in (-10^{-1} - 10^{-4}, 10^{-1} + 10^{-4})$  is much narrower. Consequently, we need fewer measurements, each of which, however, takes longer. Finally although calculating  $E - \mathcal{E}_{10^{-4}}$  is very demanding, only a very few measurements are needed since its distribution is very narrow. In the present example its width of order  $10^{-4}$  or less might even be smaller than the statistical uncertainty of  $\langle \mathcal{E}_{10^{-1}} \rangle$  such that the systematic impact on the uncertainty of  $\langle E \rangle$  would be negligible.

In the same way we can calculate averages  $\langle O(\mathcal{E}_{\delta_1}, \dots) \rangle$ ,  $\langle O(\mathcal{E}_{\delta_2}, \dots) - O(\mathcal{E}_{\delta_1}, \dots) \rangle$ , etc., although suitable values for  $\delta_1, \delta_2, \dots$  might be different for different observables. This method is a natural extension of our simulation method, and several bits of previously developed computer code could be reused for its implementation.

### E. Measuring the second virial coefficient

At the Boyle temperature  $T_B$  the second virial coefficient  $A_2$  vanishes,

$$A_2(T_B(N), N) = 0, \quad (19)$$

and the  $\Theta$  point with predicted tricritical behavior is the Boyle temperature of the infinite chain,

$$\Theta = \lim_{N \rightarrow \infty} T_B(N). \quad (20)$$

Therefore, the Boyle temperature is to be understood as the analog of the Boyle point of a real gas, i.e., the temperature where the gas is described by the ideal gas relation  $p/nRT = 1/V$ , except for higher-order corrections (third order and higher) in  $1/V$  [3]. To determine the temperature  $\Theta$ , precise measurements of the second virial coefficient  $A_2$  are essential. The latter can be expressed as

$$A_2(T, N) = \frac{\int e^{-\frac{E(\mathbf{X})+E(\mathbf{Y})}{k_B T}} \left(1 - e^{-\frac{E_{\text{inter}}(\mathbf{X}, \mathbf{Y})}{k_B T}}\right) d\mathbf{X} d\mathbf{Y}}{2Z(N, T)^2}, \quad (21)$$

where  $E_{\text{inter}}$  is the interaction between two polymer chains,

$$E_{\text{inter}}(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^N \sum_{j=1}^N U(|\mathbf{x}_i - \mathbf{y}_j|), \quad (22)$$

and  $Z$  the partition function of an isolated chain,

$$Z(N, T) = \int e^{-\frac{E(\mathbf{X})}{k_B T}} d\mathbf{X}. \quad (23)$$

Using the distance between the two center of masses

$$r_{\text{cm}}(\mathbf{X}, \mathbf{Y}) = \frac{1}{N} \left| \sum_{i=1}^N (\mathbf{x}_i - \mathbf{y}_i) \right|, \quad (24)$$

one can also write

$$A_2(T, N) = 2\pi \int_0^\infty r_{\text{cm}}^2 \langle 1 - e^{-\frac{E_{\text{inter}}}{k_B T}} \rangle_{r_{\text{cm}}} dr_{\text{cm}}, \quad (25)$$

where the average is taken for fixed values of  $r_{\text{cm}}$ .

The most commonly applied method for measuring  $A_2$  simulates two independent chains  $\mathbf{X}$  and  $\mathbf{Y}$  at temperature  $T$  to generate independent pairs of configurations  $(\mathbf{X}_t, \mathbf{Y}_t)$ . For each pair  $E_{\text{inter}}$  is sampled by placing the two chains at different distances and orientations with respect to each other and measuring the interaction. This can be done specifically by first shifting the two polymers such that their centers of gravity reside at the origin. Then a random rotation is applied to one of them, and finally one chain is shifted to a random point within a sphere around the origin. The radius of the sphere has to be large enough to include all distances that allow relevant interaction between the two polymers. It is easy to see that the more frequent occurrence of large distances renders this simple method somewhat inefficient, and an obvious improvement would be the introduction of a weight function and consequently a weighted average to enhance the selection of smaller distances. Regardless of this issue, the main problem with this strategy is that there are rare cases where the two chains interact strongly leading to a comparatively large negative  $E_{\text{inter}}$ . Since  $-E_{\text{inter}}$  appears in the exponent in (25) these cases have an extremely high

impact, but for longer chains they are practically impossible to find with the described method since as far as the measurement of  $E_{\text{inter}}$  is concerned it performs only simple sampling. For this reason we apply this method only for polymers with  $N \leq 32$ .

A method for investigating longer chains in the case of interacting self-avoiding lattice walks has been provided by Grassberger and Hegger [9] in the form of importance sampling. They propose to simulate an ensemble where the two chains interact and microstates occur with a probability proportional to the weight function

$$W_{\text{GH}}(\mathbf{X}, \mathbf{Y}) = e^{-\frac{E(\mathbf{X})}{k_B T}} e^{-\frac{E(\mathbf{Y})}{k_B T}} e^{-\frac{E_{\text{inter}}(\mathbf{X}, \mathbf{Y})}{k_B T}}. \quad (26)$$

Now the states with negative  $E_{\text{inter}}$  are sampled much more frequently, and through reweighting they contribute to the average (25) with equal impact. We choose to modify this weight function for the model considered here. The contributions from the individual chains remain

$$W(\mathbf{X}, \mathbf{Y}) = e^{-\frac{E(\mathbf{X})}{k_B T}} e^{-\frac{E(\mathbf{Y})}{k_B T}} W_{\text{inter}}(\mathbf{X}, \mathbf{Y}), \quad (27)$$

and for the contribution of the interchain interaction we use in the case of the truncated potential

$$W_{\text{inter}}(\mathbf{X}, \mathbf{Y}) = \begin{cases} e^{-\frac{E_{\text{inter}}(\mathbf{X}, \mathbf{Y})}{k_B T}} & \text{if } E_{\text{inter}} < -10^{-6}\epsilon, \\ 1.0 & \text{if } E_{\text{inter}} > 10^{-6}\epsilon, \\ e^{-r_{\text{cm}}(\mathbf{X}, \mathbf{Y})/\rho} & \text{otherwise,} \end{cases} \quad (28)$$

where  $r_{\text{cm}}$  is again the distance between the two center of masses. Now the regions of state space with positive chain-chain interaction—which do not exist for Grassberger and Hegger's model—are not suppressed and do not form barriers hampering the simulation. To prevent the two chains from permanently losing contact, the distance between the two chains is now constrained by the term  $\exp[-r_{\text{cm}}(\mathbf{X}, \mathbf{Y})/\rho]$  in the weight function, if the interchain interaction is close to zero. Here  $\rho \propto \sqrt{N}$  is a scale factor.

The case of nontruncated potential is more complicated. As with the energy of the individual chains, it is highly desirable to avoid calculating or updating the full chain-chain interaction with each MC move due to the high computational cost. Therefore, for generating the weight  $W_{\text{inter}}$  we opted to use  $\alpha E_{\text{inter}}^c$  as a proxy for  $E_{\text{inter}}$ . Here  $E_{\text{inter}}^c$  is the interchain interaction with a cutoff radius of  $2b \approx 2.2\sigma$ , and  $\alpha$  is a factor adjusted such that at the temperatures simulated  $\max\{\exp[-(E_{\text{inter}} - \alpha E_{\text{inter}}^c)/k_B T]\} \approx 1$ . The actual measurements are, of course, taken for the correct quantity  $E_{\text{inter}}$ , are appropriately reweighted, and are accelerated by the means described in the previous subsection.

## IV. RESULTS AND DISCUSSION

### A. Logarithmic corrections and $N_0$

In his theoretical analysis of the  $\Theta$  transition, Duplantier [5–7] modeled the polymer as a continuous path of length  $S$  and introduced a ultraviolet cutoff  $s_0$  marking the minimal distance along the path for which the polymer can interact with itself. Then through  $S/s_0 \rightarrow \infty$  the thermodynamic limit is approached. As a consequence scaling relations are expressed as functions of  $S/s_0$ , and while  $s_0$  can be

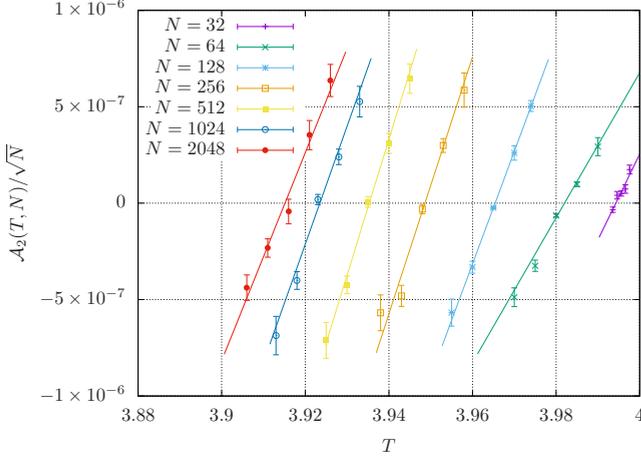


FIG. 2. Measured values for  $\mathcal{A}_2(T, N)$  and linear fits for the untruncated potential ( $r_c = \infty$ ).

absorbed by the prefactor of algebraic terms  $a(S/s_0)^x = a'S^x$ , its appearance in terms of logarithmic corrections cannot be avoided.

Universality dictates that asymptotically the results of Duplantier must also be correct for the model we consider here if we replace  $S/s_0$  by  $N/N_0$ , which raises the questions of the value of the constant  $N_0$ . Naively, the fact that interacting monomers must be separated by at least two bonds suggests  $N_0 = 2$ . However, in an imagined renormalization process one would expect this value to be subject to change. Unfortunately, the data we obtain are not precise enough and the chains we can investigate are not long enough to reliably determine  $N_0$  through fits. One would expect  $N_0$  to remain of the order of one, and this seems to agree with our results.

### B. Determination of the transition temperature $\Theta$

We performed simulations for four different cutoff radii ( $r_c/\sigma \in \{2, 3, 4, 5\}$ ) and the untruncated potential ( $r_c = \infty$ ). We investigated chain lengths  $N \in \{2, 3, \dots, 32\}$  using the simple-sampling method and  $N \in \{8 \times 2^k, 12 \times 2^k, 19 \times 2^k\}$  with integer  $k \geq 0$  using importance sampling. For  $r_c = 2\sigma$  it is  $N \leq 8192$  and for the other cases  $N \leq 4096$ .

For the determination of the Boyle temperature for each system we selected five temperatures in a relatively small interval that includes  $T_B$  and measure

$$\mathcal{A}_2(T, N) = \left\langle \frac{1 - e^{-\frac{E_{\text{inter}}(\mathbf{X}, \mathbf{Y})}{k_B T}}}{W_{\text{inter}}(\mathbf{X}, \mathbf{Y})} \right\rangle \quad (29)$$

in the ensemble described above. Technically, due to a difference in normalization,  $\mathcal{A}_2(T, N) \neq A_2(T, N)$  but  $\mathcal{A}_2(T, N) = 0$  does imply  $A_2(T, N) = 0$ . As can be seen in Fig. 2 using small temperature intervals has two advantages. On the one hand the dependency on  $T$  is approximately linear and  $T_B$  can be determined without reference to theoretical predictions, while on the other hand it appears that for simulations of equal length the statistical uncertainty of  $\mathcal{A}_2$  becomes smallest in the proximity of  $T_B$  where  $\mathcal{A}_2 \approx 0$ . We obtained estimates of the statistical errors of  $T_B$  via the fitting process, and due to

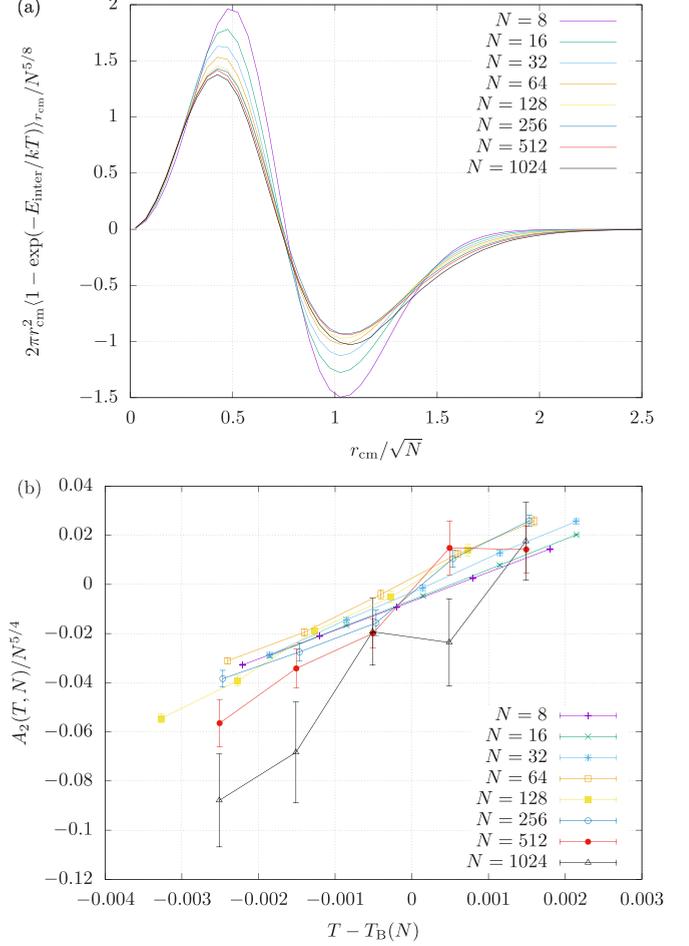


FIG. 3. (a) Rescaled contributions to the second virial coefficient for different sizes  $N$  at  $T_B(N)$ . (b) Second virial coefficient  $A_2$  in the proximity of the Boyle temperature  $T_B(N)$ . [The slight deviation from  $A_2(T_B(N), N) = 0$  is caused by systematic binning errors.] For both plots  $r_c = 2\sigma$ . Both exponents of  $N$  in the ordinates have been chosen merely for convenience and do not signify suspected asymptotic behavior.

the relatively small number of data points they themselves are relatively uncertain.

For systems with truncated potential where  $E_{\text{inter}}$  is measured exactly every time we also determined the distance-dependent averages used in Eq. (25), which are shown in Fig. 3(a). In principle these allow the calculation of the true second virial coefficient  $A_2(T, N)$  [Fig. 3(b)]. However, even with truncated potential we used  $\mathcal{A}_2(T, N)$  to determine  $T_B$  since the binning process needed to obtain  $A_2(T, N)$  this way can introduce systematic errors.

The Boyle temperatures we obtained through linear fits of  $\mathcal{A}_2$  are displayed in Fig. 4. As expected the introduction and reduction of the cutoff radius reduce the transition temperature substantially. Even a relatively distant cutoff at  $r_c = 5\sigma$  where  $U(r_c) \approx -\varepsilon/400$  changes  $T_B(N)$  for medium and large  $N$  by more than 5%.

According to renormalization-group calculations [5], for long chains one should find  $\Theta - T_B \propto N^{-1/2} (\ln N/N_0)^{-7/11}$ , where we would expect  $N_0 \approx 1$ . Therefore, we fitted the

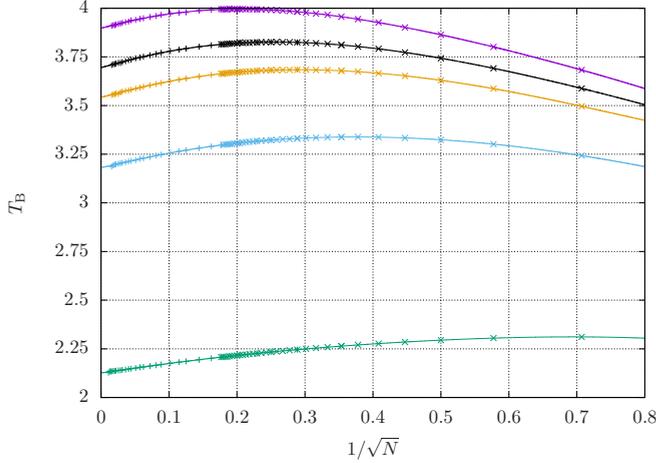


FIG. 4. Boyle temperatures for chains of different lengths  $N$ . From top to bottom the cutoff radius  $r_c$  equals  $\infty$ ,  $5\sigma$ ,  $4\sigma$ ,  $3\sigma$ , and  $2\sigma$ . Values are obtained by simple sampling ( $\times$ ) and importance sampling ( $+$ ). Lines show the fitted fourth-order polynomials with logarithmic corrections  $q_l(N)$ .

Ansatz

$$q_l(N) = \Theta_l + \frac{a_1}{N^{1/2}(\ln N/N_0)^{7/11}} + \frac{a_2}{N^{2/2}} + \frac{a_3}{N^{3/2}} + \frac{a_4}{N^{4/2}} \quad (30)$$

and found even very good agreement for all five data sets. The nature of (30) dictates that  $N_0$  has to be smaller than the smallest value of  $N$  in the data set, but beyond this condition its precise value seems of minor importance. For  $N \geq 2$  we obtain very good fits for a wide range of values with  $N_0 \lesssim 0.8$ , and the values of  $\Theta_l$  exhibit only minor variation of an order of  $10^{-3}$ . If data for small  $N$  are excluded from the fit, larger values for  $N_0$  lead to good fits as well.

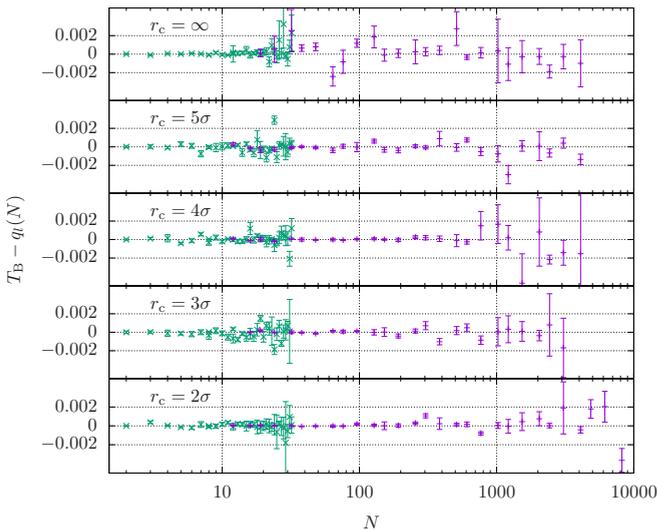


FIG. 5. Difference between measured Boyle temperatures and fitted functions  $q_l$ . Values are obtained by simple sampling ( $\times$ ) and importance sampling ( $+$ ).

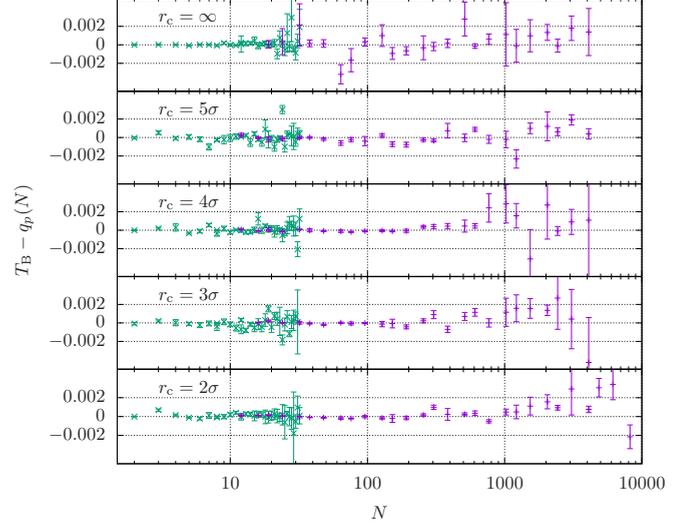


FIG. 6. Same as Fig. 5 for  $q_p$ .

It is not clear *a priori* whether logarithmic corrections are actually traceable for the model we investigate with the system sizes considered here. To test this we also fitted the bare polynomial

$$q_p(N) = \Theta_p + \frac{b_1}{N^{1/2}} + \frac{b_2}{N^{2/2}} + \frac{b_3}{N^{3/2}} + \frac{b_4}{N^{4/2}}, \quad (31)$$

which also modeled the data very well albeit not as well as (30). Irrespective of the value of  $r_c$  we find  $\Theta_p < \Theta_l$ .

Finally we also fitted a polynomial of fourth order to the inverse of the Boyle temperature assuming the relation

$$q_i(N) = \left( \frac{1}{\Theta_i} + \frac{c_1}{N^{1/2}} + \frac{c_2}{N^{2/2}} + \frac{c_3}{N^{3/2}} + \frac{c_4}{N^{4/2}} \right)^{-1}, \quad (32)$$

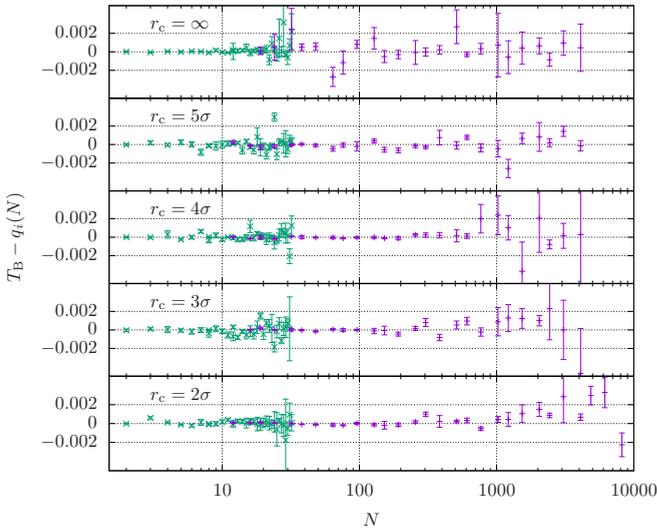
which has to a lesser degree also been employed in the past [25]. Values for  $\Theta_i$  are slightly larger but very similar to  $\Theta_p$ .

We compare all estimates for the  $\Theta$  temperatures in Table I. Although the choice of the higher-order correction terms used in Eq. (30) can certainly be debated, at the very least we get an impression on how weakly the estimate of the  $\Theta$  temperature from our data is influenced by the assumed analytical function.

All three functions (30), (31), and (32) model the data very well and would be indistinguishable in a plot like Fig. 4. In order to “zoom in” and to better appreciate the (dis-)agreement, the differences between the data and the fits are shown in Figs. 5–7. One can attempt to judge whether the data thus

TABLE I.  $\Theta$  temperatures obtained through fits of Eqs. (31), (32), and (30).

$r_c$	$\Theta_p$	$\Theta_i$	$\Theta_l$
$\infty$	3.8921(3)	3.8938(3)	3.8979(5)
$5\sigma$	3.6907(4)	3.6917(3)	3.6954(9)
$4\sigma$	3.5368(2)	3.5380(2)	3.5427(3)
$3\sigma$	3.1761(3)	3.1768(3)	3.1814(4)
$2\sigma$	2.1236(2)	2.1238(2)	2.1265(2)


 FIG. 7. Same as Fig. 5 for  $q_i$ .

displayed show merely statistical noise by randomly fluctuating around the zero value or whether they constitute a trend that indicates systematic deviations. It seems that fits of  $q_p$  and  $q_i$  both underestimate the Boyle temperatures for long chains and that  $q_l$  represents the behavior best. For  $r_c = 2\sigma$  the longest chains could be simulated, and we fitted in Fig. 8 the three functional forms to the subset of smaller chains  $N < 400$  and used them as predictive models for the longer ones. While  $q_p$  and  $q_i$  now clearly underestimate the Boyle temperature,  $q_l$  with the logarithmic corrections performs much better. Our results thus provide evidence for nonpolynomial scaling; specific details, e.g., the exponent  $7/11$ , unfortunately, cannot be verified at this point.

Since the  $\Theta$  point can also be determined by geometric means, the value obtained through extrapolation of the Boyle temperatures can be tested. The method we use is similar to the technique employed in [12]. At the transition the polymer

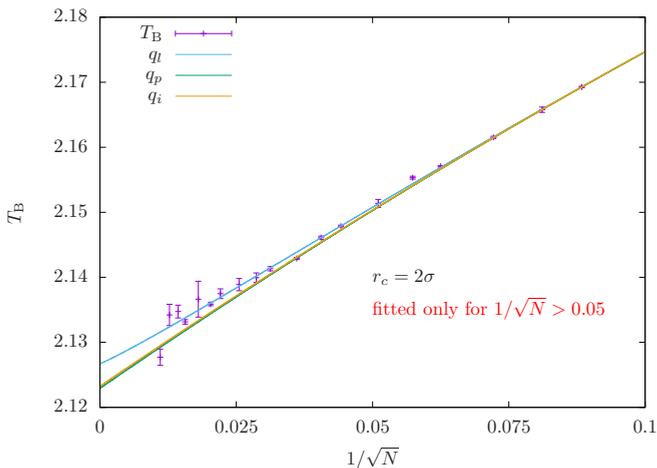


FIG. 8. If the functions  $q_l$ ,  $q_p$ , and  $q_i$  are fitted using only data for  $N < 400$ , the ansatz  $q_l$ , which includes logarithmic corrections, best predicts the behavior of long chains, whereas  $q_p$  and  $q_i$  (almost identical curves) generate too small values for  $T_B$ .

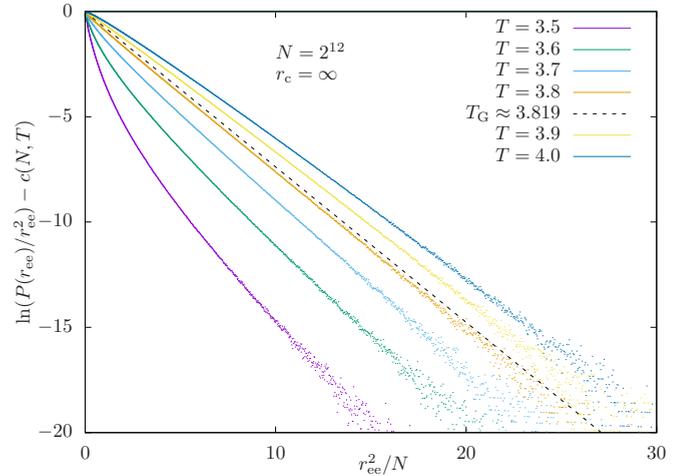


FIG. 9. Fitting quadratic polynomials and interpolating the fitted coefficients of the linear and quadratic terms as functions of  $T$  allows one to estimate the temperature  $T_G$  where the relationship (34) is fulfilled. There the second-order term vanishes, and the remaining linear term leads to the dotted line. The polynomial's constant term  $c(N, T)$  has been subtracted to allow easier visual comparison.

behaves at large enough scales like a Gaussian random walk, which implies that the end-to-end vector should be distributed according to

$$P(\mathbf{r}_{ee}) \propto e^{-r_{ee}^2/2\sigma_{r_{ee}}(N, T)^2}. \quad (33)$$

This means that for the distribution of the end-to-end distance  $r_{ee} = |\mathbf{r}_{ee}|$  we find

$$\ln(P(r_{ee})/r_{ee}^2) = -r_{ee}^2/2\sigma_{r_{ee}}(N, T)^2 + c(N, T), \quad (34)$$

where  $c(N, T)$  is constant with respect to  $r_{ee}$ . We, therefore, can define a ‘‘Gauss’’ temperature  $T_G$  as the temperature where the second-order term of a quadratic polynomial fitted to  $\ln(P(r_{ee})/r_{ee}^2)$  as a function of  $r_{ee}^2$  vanishes (see Fig. 9).

We measured  $T_G$  for  $r_c = \infty$  and various chain lengths and show the results together with the Boyle temperature in Fig. 10. For finite  $N$  we find  $T_G < T_B$ . It is apparent that in the thermodynamic limit the different temperatures converge or become at least very close, but also that for any chain length  $T_G$  differs from  $\Theta$  much more strongly than  $T_B$  does. Considering further that we do not know which scaling law  $T_G$  is supposed to follow, we have to conclude that this geometric method of determining  $\Theta$  is on its own less suitable than the second-virial-coefficient technique even though larger systems can be treated.

### C. The polymer at the $\Theta$ temperature

For the investigation of the polymer in the tricritical region we use the temperature obtained with logarithmic corrections taken into account:  $\Theta \equiv \Theta_l$ .

We performed simulations close to  $\Theta$  for  $r_c = \infty$  and  $r_c = 2\sigma$  to investigate scaling behavior. Here chains with up to  $N = 32\,768$  are considered. It is well known and has been verified numerous times that the size of the chain measured by  $\sqrt{\langle r_{ee}^2 \rangle}$  or  $\sqrt{\langle r_{gyr}^2 \rangle}$  asymptotically scales like  $\sqrt{N}$  as is expected from a Gaussian random walk. This means that for

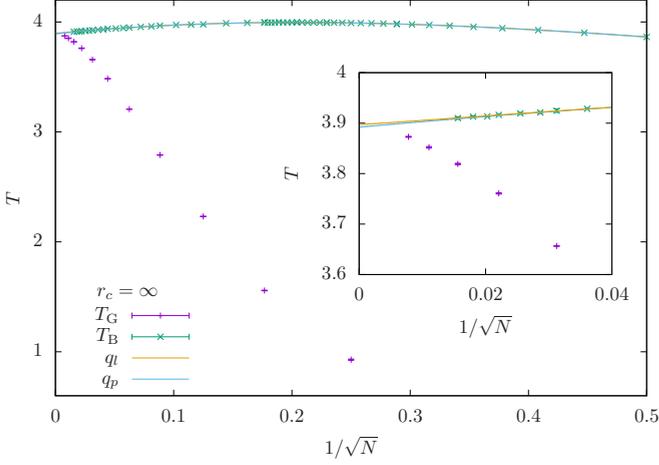


FIG. 10. Comparing  $T_G$  to the Boyle temperatures shows good agreement in the thermodynamic limit. The inset magnifies details of the main plot for  $N > 600$ .

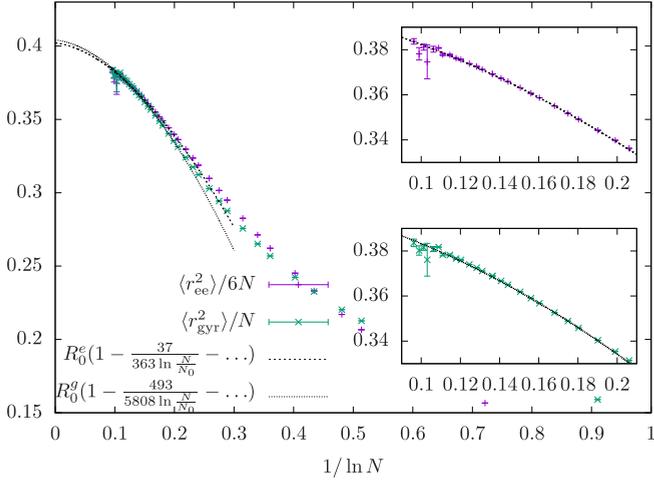


FIG. 11. Squared radius of gyration and end-to-end distance at the  $\Theta$  temperature ( $r_c = \infty$ ). The insets show for  $N > 100$  the same data sets individually with their respective fits.

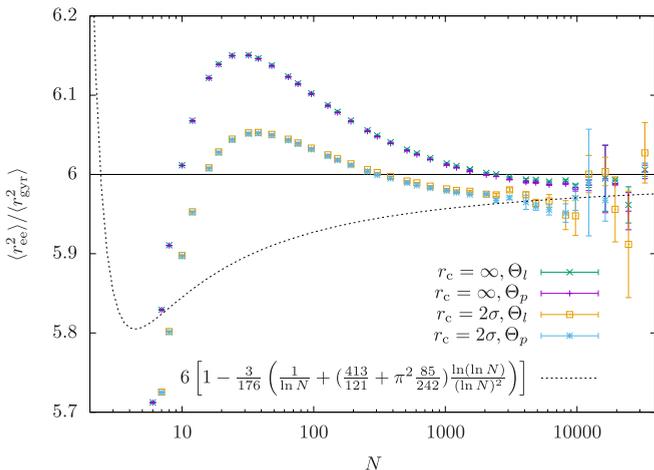


FIG. 12. Ratio of squared end-to-end distance and squared radius of gyration at the  $\Theta$  temperature and the analytical prediction.

$\langle r_X^2 \rangle / N$  a nonzero limit for  $N \rightarrow \infty$  exists. Unfortunately, logarithmic corrections which are expected here are generally difficult to verify. For the case of the tricritical  $\Theta$  transition in particular Hager and Schäfer [8] have shown that at least first-order ( $\propto 1/\ln N$ ) and second-order [ $\propto \ln(\ln N)/(\ln N)^2$ ] corrections are of the same magnitude for the chain length considered here, and both have to be included in the analysis. They found that

$$\frac{\langle r_{ee}^2 \rangle}{N} = 6R_0 \left( 1 - \frac{37}{363 \ln \frac{N}{N_0}} - \frac{37\kappa \ln \left( \ln \frac{N}{N_0} \right)}{363 \ln^2 \frac{N}{N_0}} + O((\ln N)^{-2}) \right), \quad (35)$$

$$\frac{\langle r_{gyr}^2 \rangle}{N} = R_0 \left( 1 - \frac{493}{5808 \ln \frac{N}{N_0}} - \frac{493\kappa \ln \left( \ln \frac{N}{N_0} \right)}{5808 \ln^2 \frac{N}{N_0}} + O((\ln N)^{-2}) \right), \quad (36)$$

and consequently

$$\frac{\langle r_{ee}^2 \rangle}{\langle r_{gyr}^2 \rangle} = 6 \left( 1 - \frac{3}{176 \ln \frac{N}{N_0}} - \frac{3\kappa \ln \left( \ln \frac{N}{N_0} \right)}{176 \ln^2 \frac{N}{N_0}} + O((\ln N)^{-2}) \right), \quad (37)$$

where  $\kappa = \frac{413}{212} + \pi^2 \frac{85}{242} \approx 6.88$  is a numerical constant. In Fig. 11 we attempt to fit these relations to  $\langle r_{ee}^2 \rangle / 6N$  and  $\langle r_{gyr}^2 \rangle / N$  for  $(\ln N)^{-1} < 0.2$ , i.e., for  $N > 148$ . We included a term  $\propto (\ln N)^{-2}$ . If we set  $N_0 \approx 1$ , the individual fits achieve rather good agreement with the data albeit not within error bars. Also, the respective estimates for  $R_0$  differ by about 0.5%:  $R_0^e \approx 0.402$  vs  $R_0^g \approx 0.404$ . Although our data are comparatively precise, we are still unable to verify or reject the predictions. We can at least note that our results are not in conflict with theory, and we find no reason to doubt that data for larger chains would follow Eqs. (35) and (36) more closely converging to the same limit  $R_0$ . In Fig. 12 we show the ratio  $\langle r_{ee}^2 \rangle / \langle r_{gyr}^2 \rangle$  together with the prediction (37) as a function of  $\ln N$ . For the latter the argument should actually be  $N/N_0$  instead of  $N$  with an unknown constant  $N_0$  leading to a horizontal shift of the curve. Regardless of this issue, the statistical errors are too large and the sizes still too small to draw definite conclusion. It is clear that asymptotic behavior can be observed only for long chains, and our data show that for this model precise data with at least  $N > 10\,000$  are needed. We show data for both  $\Theta_l$  and  $\Theta_p$  to illustrate that the uncertainty about the true value of  $\Theta$  plays a very minor role when it comes to the agreement or disagreement with theoretical predictions.

For the temperature derivative of the squared radius of gyration and end-to-end distance the calculations of Duplantier [7] predict

$$\frac{\partial}{\partial T} \ln \langle r_X^2 \rangle \sim \frac{\sqrt{N}}{(\ln N / N_0)^{4/11}} \quad (38)$$

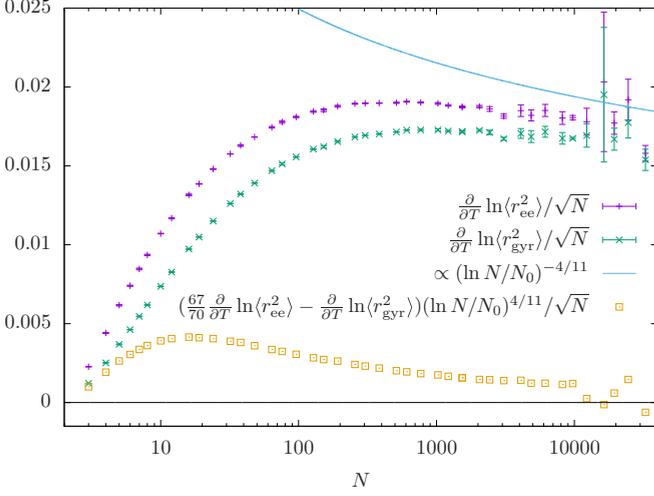


FIG. 13. Temperature derivative of the logarithmic squared radius of gyration and end-to-end distance at the  $\Theta$  temperature and the predicted asymptotic scaling ( $r_c = \infty$ ).

and

$$\frac{\frac{\partial}{\partial T} \ln \langle r_{ee}^2 \rangle}{\frac{\partial}{\partial T} \ln \langle r_{gyr}^2 \rangle} = \frac{70}{67} \quad (39)$$

for long chains. We took derivatives of quadratic polynomials fitted to the canonical averages at temperatures close to the  $\Theta$  temperature:  $|\Theta - T| \leq 0.02$ . Using linear polynomials instead produced almost identical results, which suggests that this procedure is valid. In Fig. 13 we display the derivatives of the logarithmic averages along with the rescaled difference

$$\left( \frac{67}{70} \frac{\partial}{\partial T} \ln \langle r_{ee}^2 \rangle - \frac{\partial}{\partial T} \ln \langle r_{gyr}^2 \rangle \right) \left( \ln \frac{N}{N_0} \right)^{4/11} / \sqrt{N} \quad (40)$$

setting  $N_0 = 1$ . For the latter we are unable to estimate statistical uncertainties since the data for  $\langle r_{ee}^2 \rangle$  and  $\langle r_{gyr}^2 \rangle$  are highly

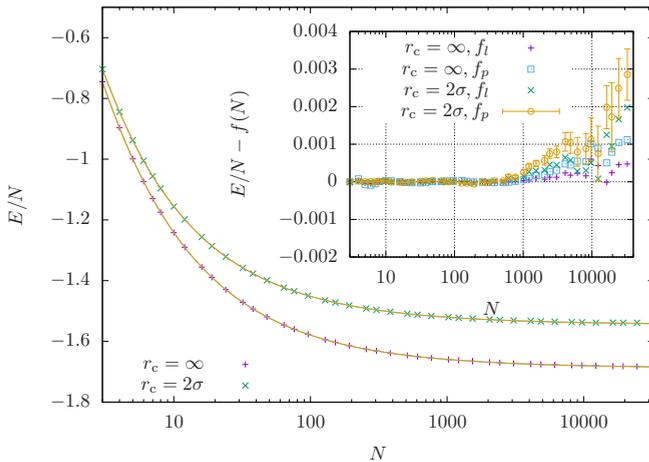


FIG. 14. Energy per monomer as function of system size at the  $\Theta$  temperature for  $r_c = \infty$  and  $r_c = 2\sigma$ . The inset shows the deviation of the data from the fitted functions  $f_l$  and  $f_p$  from Eqs. (41) and (42), respectively. Error bars are displayed exemplarily only for one data set to preserve visibility of all data points.

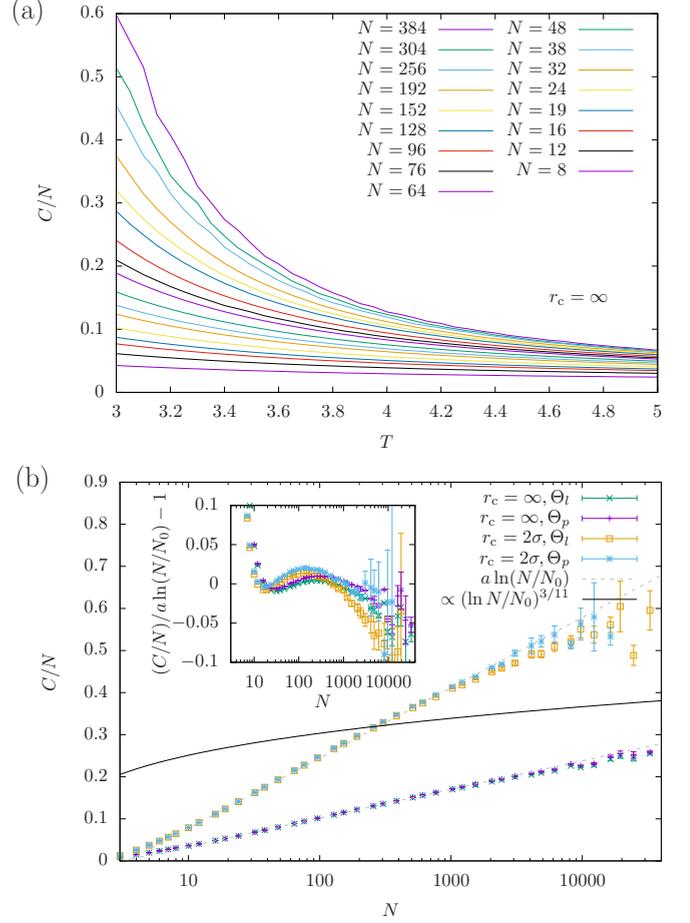


FIG. 15. (a) Specific heat per monomer  $C/N$  as function of temperature  $T$  for different sizes  $N$ . (b) Specific heat per monomer  $C/N$  at the  $\Theta$  temperature for different sizes  $N$ . As a visual reference we included the predicted asymptotic scaling using  $N_0 = 1$  (black line). Although the first impression is that  $C/N$  scales linearly with  $\ln(N/N_0)$  (dashed lines), a closer inspection (inset) reveals systematic deviations which increase with system size.  $a$  is a fitting parameter. The symbols in the inset belong to the same parameter combinations as in the main plot.

correlated. We find that for long enough chains  $\frac{\partial}{\partial T} \ln \langle r_x^2 \rangle / \sqrt{N}$  varies only very little, even though a decrease for  $N > 1000$  can be observed and the data might approach the predicted asymptotic behavior. However, there is too much noise, and the chains are still too short to be certain. It also seems likely that the difference (40) approaches zero as predicted, albeit very slowly. On the one hand this observation supports the theory, while on the other it demonstrates that even the behavior of the chains we investigate here which are of substantial length is still considerably different from that of the infinite chain.

How the energy per monomer depends on  $N$  at  $\Theta_l$  is plotted in Fig. 14. We fitted the two functions

$$f_l(N) = e_l + \frac{\beta_1}{N^{1/2} [\ln(N/N_0)]^{4/11}} + \frac{\beta_2}{N^{2/2}} + \dots, \quad (41)$$

$$f_p(N) = e_p + \frac{\alpha_1}{N^{1/2}} + \frac{\alpha_2}{N^{2/2}} + \dots \quad (42)$$

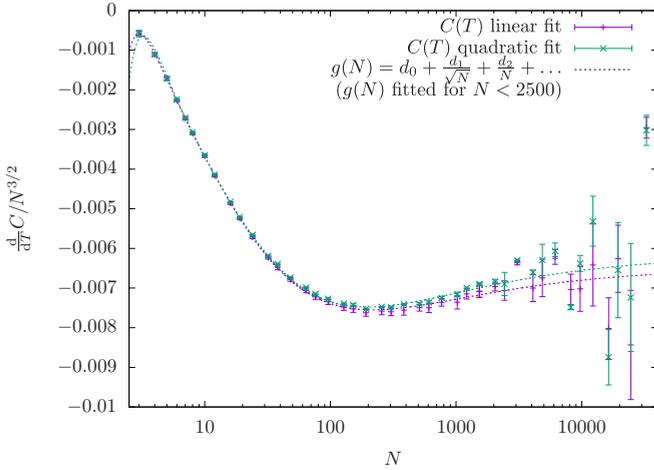


FIG. 16. Temperature derivative of the specific heat  $dC/dT$  at the  $\Theta$  temperature for different chain lengths  $N$ .

to the data with the predicted asymptotic behavior given by the first two terms of  $f_i$  [2]. Again, the value of  $N_0$  is restricted by the smallest  $N$  in the data, and it does not change the quality of the fit much. In any case, the functional form with logarithmic corrections fits the data better than the bare polynomial. For Fig. 14 all data are included ( $N \geq 3$ ) and  $N_0 = 0.3$ . With these conditions and  $r_c = \infty$  we obtain  $e_l = -1.68774(4)$  and  $e_p = -1.6890(1)$ , while for  $r_c = 2\sigma$  it is  $e_l = -1.5445(1)$  and  $e_p = -1.5460(2)$ .

We used the opportunity to successfully verify our results for the specific heat as function of temperature shown in Fig. 15(a) by comparing to an earlier study [26] of the same model [27]. At the  $\Theta$  temperature the specific heat is predicted [6] to behave as  $C/N \propto (\ln N/N_0)^{3/11}$ . At first sight our data in Fig. 15(b) show a very different picture, and a linear dependence on  $\ln N$ , i.e., Gaussian chain behavior, seems to be apparent. However, deviations from this behavior are present at all sizes (see inset), and the scaling finally changes at  $N \approx 10^4$  to what might very well be the predicted behavior. We also see that small changes in the temperature, e.g., using  $\Theta_p$  instead of  $\Theta_l$ , hardly change the scaling.

Finally we measured the temperature derivative of the specific heat at the  $\Theta$  temperature. This was done by fitting polynomials of first and second order to  $C(T)$  on an interval  $[3.85, 3.93]$  containing  $\Theta$  and taking their derivatives. Because for each size  $N$  only a few data points (temperatures) are available, the error estimates themselves, which were obtained through the fitting process, are relatively uncertain. The results are shown in Fig. 16. To our knowledge there is no analytical prediction for  $dC(T)/dT$  since the partition function has not been developed beyond second order. The data suggest that in the limit of long chains  $dC(T)/dT|_{T=\Theta}/N \propto N^{1/2}$ . However, a straightforward fit of a polynomial in powers of  $N^{-1/2}$  to  $dC(T)/dT|_{T=\Theta}/N^{3/2}$  does not work very well. We suspect that logarithmic corrections could play a role here too.

## V. CONCLUSION

In this study we applied the algorithm we introduced in a recent publication [17] to investigate a polymer model at the

$\Theta$  transition. Within the framework of the algorithm, decisions on proposed updates are made with as little information about the associated change in energy as possible. Since the precise evaluation of the change in energy, which is computationally very expensive, is thus avoided, the method is very fast. It is now possible to simulate chains in continuous space with tens of thousands of monomers even if the potential is not truncated and each monomer interacts with all others. The accessible chain lengths are now more than an order of magnitude larger than what has been achieved previously in similar studies [13] and are comparable to what has been done for lattice models with short-range next-neighbor interaction [9,10]. Although the model we have used here is rather simple, the algorithm is versatile, and different pair potentials, elastic bonds, stiffness, or heterogeneity can easily be incorporated.

True to the philosophy of the algorithm we developed a technique to accurately and efficiently measure expectation values of geometric and thermodynamic observables, which allows for uncertainties in the majority of individual measurements. Although the underlying idea is rather simple and it seems likely that it has been found before, we are unaware of any previous application.

We measured the second virial coefficient in order to determine the Boyle temperature and systematically investigated how the latter is influenced by a truncation of the monomer-monomer interaction potential. With the exception of a very small cutoff distance, with increasing size the Boyle temperature increases, reaches a maximum, and approaches the infinite chain limit, i.e., the  $\Theta$  temperature from above. Truncation of the potential has a big influence and reduces the Boyle temperature up to about 50%. Data for all cutoff radii can be very well modeled with polynomials of fourth order, yet when the linear term is modified with the predicted logarithmic correction the agreement is better still.

At the  $\Theta$  temperature the scaling of the average end-to-end distance and radius of gyration is not in conflict with Duplantier's predictions, but the agreement is also not as good as one might wish. It seems that the chain lengths we were able to simulate are still too small. On the other hand we do observe that the relationship between the temperature derivatives of the two quantities compares very well with theory. The specific heat again does not clearly demonstrate the predicted logarithmic asymptotic scaling. However, we were able to reach system sizes where the linear scaling that we observe for short chains breaks down and the curve bends in the right direction. Unfortunately, considering that several hundreds of CPUs ran for several months to produce the data presented here, we are not overly optimistic about the possibility of investigating much longer chains unless additional substantial improvements can be made to the methodology.

## ACKNOWLEDGMENT

The project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the Collaborative Research Centre under Grant No. 189 853 844–SFB/TRR 102 (project B04).

- [1] P.-G. De Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, NY, 1979).
- [2] J. Des Cloizeaux and G. Jannink, *Polymers in Solution: Their Modelling and Structure* (Oxford University Press on Demand, Oxford, UK, 1990).
- [3] P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, NY, 1953).
- [4] P.-G. De Gennes, Collapse of a polymer chain in poor solvents, *J. Phys. Lett.* **36**, 55 (1975).
- [5] B. Duplantier, Lagrangian tricritical theory of polymer chain solutions near the  $\Theta$ -point, *J. Phys.* **43**, 991 (1982).
- [6] B. Duplantier, Tricritical polymer chains in or below three dimensions, *Europhys. Lett.* **1**, 491 (1986).
- [7] B. Duplantier, Geometry of polymer chains near the theta-point and dimensional regularization, *J. Chem. Phys.* **86**, 4233 (1987).
- [8] J. Hager and L. Schäfer,  $\Theta$ -point behavior of diluted polymer solutions: Can one observe the universal logarithmic corrections predicted by field theory? *Phys. Rev. E* **60**, 2071 (1999).
- [9] P. Grassberger and R. Hegger, Simulations of three-dimensional  $\Theta$  polymers, *J. Chem. Phys.* **102**, 6881 (1995).
- [10] P. Grassberger, Pruned-enriched Rosenbluth method: Simulations of  $\theta$  polymers of chain length up to 1 000 000, *Phys. Rev. E* **56**, 3682 (1997).
- [11] V. I. Harismiadis and I. Szleifer, Second virial coefficients of chain molecules: A Monte Carlo study, *Mol. Phys.* **81**, 851 (1994).
- [12] C. Yong, J. H. Clarke, J. J. Freire, and M. Bishop, The theta condition for linear polymer chains in continuous space and three dimensions, *J. Chem. Phys.* **105**, 9666 (1996).
- [13] A. M. Rubio and J. J. Freire, Tricritical behavior of an off-lattice flexible polymer model: Monte Carlo calculation of second virial coefficients, *J. Chem. Phys.* **106**, 5638 (1997).
- [14] A. M. Rubio, J. J. Freire, C. Yong, and J. H. Clarke, Conformational properties of polymer chains in the theta region, *J. Chem. Phys.* **111**, 1302 (1999).
- [15] F. Rampf, K. Binder, and W. Paul, The phase diagram of a single polymer chain: New insights from a new simulation method, *J. Polym. Sci. B: Polym. Phys.* **44**, 2542 (2006).
- [16] P. Zhang, N. M. Alsaifi, and Z.-G. Wang, Revisiting the  $\Theta$  point, *Macromolecules* **53**, 10409 (2020).
- [17] S. Schnabel and W. Janke, Accelerating polymer simulation by means of tree data-structures and a parsimonious Metropolis algorithm, *Comput. Phys. Commun.* **256**, 107414 (2020).
- [18] N. Madras and A. D. Sokal, The pivot algorithm: A highly efficient Monte Carlo method for the self-avoiding walk, *J. Stat. Phys.* **50**, 109 (1988).
- [19] S. Schnabel, W. Janke, and M. Bachmann, Advanced multi-canonical Monte Carlo methods for efficient simulations of nucleation processes of polymers, *J. Comput. Phys.* **230**, 4454 (2011).
- [20] N. Clisby, Accurate estimate of the critical exponent  $\nu$  for self-avoiding walks via a fast implementation of the pivot algorithm, *Phys. Rev. Lett.* **104**, 055702 (2010).
- [21] N. Clisby, Efficient implementation of the pivot algorithm for self-avoiding walks, *J. Stat. Phys.* **140**, 349 (2010).
- [22] N. Clisby and B. Dünweg, High-precision estimate of the hydrodynamic radius for self-avoiding walks, *Phys. Rev. E* **94**, 052102 (2016).
- [23] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.* **21**, 1087 (1953).
- [24] F. Müller, H. Christiansen, S. Schnabel, and W. Janke, Fast, hierarchical, and adaptive algorithm for Metropolis Monte Carlo simulations of long-range interacting systems, *Phys. Rev. X* **13**, 031006 (2023).
- [25] T. Vogel, M. Bachmann, and W. Janke, Freezing and collapse of flexible polymers on regular lattices in three dimensions, *Phys. Rev. E* **76**, 061803 (2007).
- [26] A. Baumgärtner, Statics and dynamics of the freely jointed polymer chain with Lennard-Jones interaction, *J. Chem. Phys.* **72**, 871 (1980).
- [27] In this study the definition of the second virial coefficient for a gas of monomers was used which led to Boyle temperatures different from the values presented here.