Supplemental Material: Knots as a Topological Order Parameter for Semiflexible Polymers

Martin Marenz* and Wolfhard Janke†
Institut für Theoretische Physik, Universität Leipzig, Postfach 100 920, D-04009 Leipzig, Germany

This supplemental material provides more detailed information on the algorithms used to generate and analyze the data. We briefly describe the two used Monte Carlo method. First the extension of the one-dimensional replica exchange method to a further parameter, and secondly the combination of the one-dimensional replica exchange with the multicanonical method. The data generated by the two-dimensional replica exchange is an ideal basis for a two-dimensional weighted histogram analysis method, which is explained next. Finally, we describe how the employed knot closure is constructed and compare it with a different way to connect the two ends of the polymer.

MONTE CARLO ALGORITHMS

The microstates (conformations) of potentially knotted semiflexible polymers are rather complex. This gives rise to a very rugged free-energy landscape, severely hampering standard simulations [1]. To cope with this problem, we have implemented two sophisticated simulation methods described below, which are complementary to each other. As a useful cross-check, all simulations have been performed with both of them and we carefully checked that the results agree within statistical error bars.

Two-dimensional replica exchange

The two-dimensional replica exchange method (2D-RE) is an extension of the standard parallel-tempering method [2]. The parallel-tempering method simulates m different replicas at once, each at a different temperature $T_\mu = 1/\beta_\mu$. Every now and then, for two replicas μ and ν an exchange of their states $\{r\}$ is proposed. To ensure detailed balance, these exchanges are accepted with probability

$$p\left(\{r\}_\mu \to \{r\}_\nu\right) = \min\left(1, \exp\left(\Delta\beta\Delta E\right)\right), \qquad (1)$$

where $\Delta\beta = \beta_\mu - \beta_\nu$ is the difference in the inverse temperature and $\Delta E = E_\mu - E_\nu$ is the difference in the energy.

If the simulated system has a Hamiltonian of the form

$$H = E_0 + \kappa E_1, \qquad (2)$$

we can construct a two-dimensional replica exchange method which simultaneously probes the system in T and κ. This means that instead of simulating m temperatures at once, we now simulate m different parameter pairs $(T,\kappa)_m$. The exchange probability is then given as

$$p\left(\{r\}_\mu \to \{r\}_\nu\right) = \min\left(1, \exp\left(\Delta\beta\Delta E_0 + \Delta\left(\beta\kappa\right)\Delta E_1\right)\right), \qquad (3)$$

with $\Delta\left(\beta\kappa\right) = \beta_\mu\kappa_\mu - \beta_\nu\kappa_\nu$. In the two-dimensional parameter space, the simulation can avoid topological barriers which hinder the flux of the replicas. In principle, one is not restricted to systems with a Hamiltonian as in (2), but this form is needed when weighted histogram analysis is applied to the generated data.

MUCA+RE

The idea behind multicanonical sampling [3] is relatively simple – the canonical Boltzmann weight is replaced by an artificial one:

$$\begin{aligned} Z_{\mathrm{CAN}} &= \sum_E \Omega(E)\exp\left(-\beta E\right) \\ \to Z_{\mathrm{MUCA}} &= \sum_E \Omega(E) W(E). \end{aligned} \qquad (4)$$

In principle, the multicanonical weight $W(E)$ can have arbitrary form, but if it is chosen as the inverse of the density of states $W(E) = \Omega(E)^{-1}$, the simulation spends an equal amount of time at each energy, leading to a constant energy histogram $H(E)$. Thus, the simulation acts as a random walker in the energy landscape and samples all regions of the phase space, even if they are suppressed in the canonical ensemble. To calculate the canonical expectation value of an observable $\mathcal{O}$ one can reweight the resulting time series of N measurements to the inverse temperature β via

$$\langle O\rangle_\beta \approx \overline{O}_\beta = \frac{\sum_{i=1}^N O_i \exp(-\beta E_i) W^{-1}(E_i)}{\sum_{i=1}^N \exp(-\beta E_i) W^{-1}(E_i)}, \qquad (5)$$

where O_i and E_i denote the measurements at "time" i.

One problem remains: Since we do not know $\Omega(E)$ beforehand it is not obvious how to construct $W(E)$ such that the resulting histogram $H(E)$ becomes flat. A possible approach is a procedure where, for a given weight $W_n(E)$, one measures the energy histogram $H_n(E)$ and improves the weight iteratively via

$$W_{n+1}(E) = \frac{W_n(E)}{H_n(E)}, \qquad (6)$$

until the resulting histogram is sufficiently flat. The iteration can be initialized with $W_0(E) = 1$.

This algorithm can easily be parallelized [4]. Instead of simulating the system only once, one can use m different copies that have identical weights but start with a different seed for the random number generator and thus realize different Markov chains. After each iteration the histograms $H_n^i(E)$ of all replicas $i = 1, \ldots, m$ are summed up and the next weight is calculated from the summed histograms,

$$W_{n+1}(E) = \frac{W_n(E)}{\sum_{i=1}^m H_n^i(E)}. \tag{7}$$

This enables us to use many CPUs (in the present study up to 128) for one multicanonical simulation and drastically reduces the needed wall-clock time.

If the simulated system suffers from suppressed regions in the phase space which are not reflected in the energy distribution $p(E)$ (but rather in observables "orthogonal" to E), the multicanonical simulation can get stuck. As we have shown in this work, the knotting transitions are exactly of this kind. A stuck simulation results in histograms $H_n(E)$ which do not become flat, even after many iterations. In most cases this is visible as a large peak in $H_n(E)$ which tends to persist in the multicanonical iterations. To cope with this problem we added a replica exchange in "orthogonal" direction to the multicanonical method (MUCA+RE). As in the parallel-tempering method one simulates m different replicas, in our case each at a different parameter κ_k, $k = 1, \ldots, m$. The underlying method is a multicanonical simulation using weights $W_\kappa(E)$ instead of a Metropolis simulation with weights $e^{-\beta_k(E_0+\kappa_k E_1)}$. The exchange probability thus becomes

$$p\left(\{r\}_\mu \to \{r\}_\nu\right) = \min\left(1, \frac{W_\kappa(E_\mu^{\kappa'}) W_{\kappa'}(E_\nu^\kappa)}{W_{\kappa'}(E_\nu^{\kappa'}) W_\kappa(E_\mu^\kappa)}\right), \tag{8}$$

where $\{r\}_\mu$ is the state of the replica at κ with weight W_κ and $\{r\}_\nu$ is the state of the other replica at κ' with weight $W_{\kappa'}$. The term E_μ^κ denotes the energy of state $\{r\}_\mu$ for the Hamiltonian defined by κ.

TWO-DIMENSIONAL WEIGHTED HISTOGRAM ANALYSIS METHOD

The one-dimensional weighted histogram analysis method (WHAM) [5] was originally developed for an optimal combination of data from several canonical simulations, each at a different inverse temperature β_i. By employing the WHAM equations [5]

$$\Omega(E) = \frac{\sum_{k=1}^m g_k^{-1} H_k(E)}{\sum_{k=1}^m N_k g_k^{-1} Z_{\beta_k}^{-1} e^{-\beta_k E}}, \tag{9}$$

$$\begin{aligned} Z_{\beta_i} &= \sum_E \Omega(E) e^{-\beta_i E} \\ &= \sum_E e^{-\beta_i E} \frac{\sum_{k=1}^m g_k^{-1} H_k(E)}{\sum_{k=1}^m N_k g_k^{-1} Z_{\beta_k}^{-1} e^{-\beta_k E}}, \end{aligned} \tag{10}$$

one can self-consistently calculate the density of states $\Omega(E)$. One starts with an initial guess of the (a priori) unknown Z_{β_i} and iterates Eq. (10) until the Z_{β_i} reach a fixed point. Then Eq. (9) can be used to calculate $\Omega(E)$. In Eqs. (9) and (10), $H_k(E)$ are the measured histograms at β_k, N_k is the number of data points used to generate $H_k(E)$, and $g_k = 1 + 2\tau_k$ accounts for the integrated autocorrelation time τ_k.

With the data obtained from a two-dimensional replica exchange method on a system with a Hamiltonian of form (2), one can easily extend the WHAM equations by an additional parameter:

$$\Omega(E_0, E_1) = \frac{\sum_{k=1}^m g_k^{-1} H_k(E_0, E_1)}{\sum_{k=1}^m N_k g_k^{-1} Z_{\beta_k,\kappa_k}^{-1} e^{-\beta_k(E_0+\kappa_k E_1)}}, \tag{11}$$

$$\begin{aligned} Z_{\beta_i,\kappa_i} = \sum_{E_0,E_1} \Bigg[& e^{-\beta_i(E_0+\kappa_i E_1)} \times \\ & \frac{\sum_{k=1}^m g_k^{-1} H_k(E_0, E_1)}{\sum_{k=1}^m N_k g_k^{-1} Z_{\beta_k,\kappa_k}^{-1} e^{-\beta_k(E_0+\kappa_k E_1)}} \Bigg]. \end{aligned} \tag{12}$$

$H_k(E_0, E_1)$ is now the two-dimensional histogram measured at $(\beta, \kappa)_k$. Having determined the two-dimensional density of states $\Omega(E_0, E_1)$, the canonical expectation value of an observable $\mathcal{O}$ at any β and κ can be calculated similarly to (5),

$$\overline{O}_{\beta,\kappa} = \frac{\sum_{E_0,E_1} O(E_0, E_1) \Omega(E_0, E_1) e^{-\beta(E_0+\kappa E_1)}}{\sum_{E_0,E_1} \Omega(E_0, E_1) e^{-\beta(E_0+\kappa E_1)}}. \tag{13}$$

KNOT CLOSURES

To identify knots in open polymers one has to apply a closure prescription. The easiest one would be to draw a virtual bond connecting the two termini of the polymer, but this direct closure would result in quite complicated knots when the polymer is very compact. In this work we employed the closure CI, which is inspired by tying a real knot, see Fig. 1. First we connect both termini of the polymer by a straight line. This connecting line is then extended in both directions, so that we get two new virtual points A' and B' located far away from all monomers. We create a third virtual point C on a perpendicular bisector of the connecting line and also far away from all

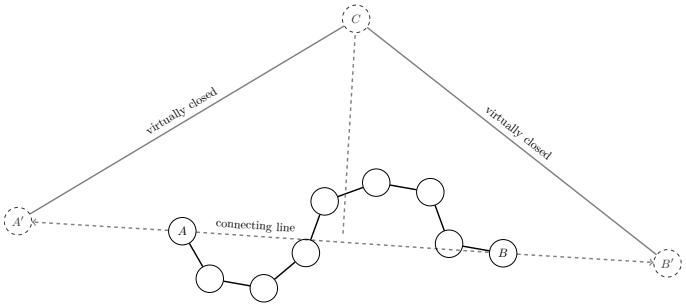

FIG. 1. The closure CI shifts the two termini of the polymer, A and B, on the line connecting these termini far outside the polymer to points A' and B'. A new point C is created on a line perpendicular to the connecting line and the points A' and B' are virtually closed through C.

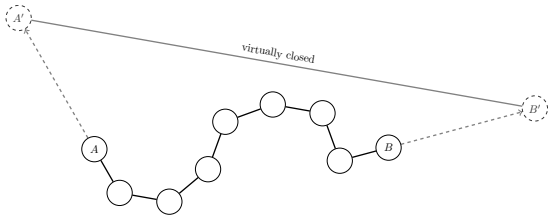

FIG. 2. The closure CII shifts the two termini, A and B, to points far outside the polymer and then closes the new termini, A' and B', virtually.

monomers. The polymer is now closed via straight lines connecting C with A' and B', respectively. For testing purpose we also implemented a second closure, CII. Here one virtually increases the bond length of the first and last bond to two new virtual termini A' and B' that are then connected by a straight line to close the polymer, see Fig. 2.

Figure 3 shows that the closures CI and CII give qualitatively similar results, and are both suitable for identifying the knotted region. However, with CI all measured states in the low-temperature region have the identical knot type $K3_1$ which means an identical value of the topological order parameter $D = 9.05463\ldots$, and thus, no statistical fluctuations are observable.

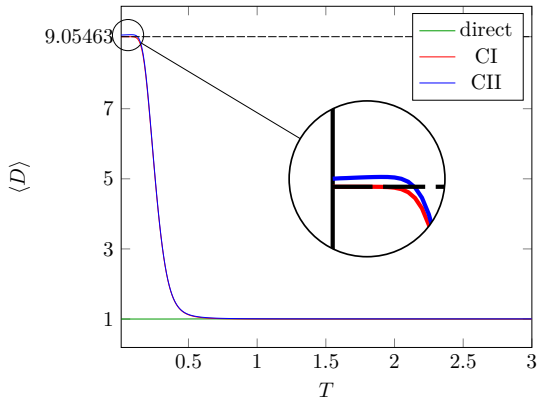

FIG. 3. Topological order parameter $\langle D\rangle$ measured by employing the closures CI, CII, and a direct closure for a 14mer at $\kappa = 3.0$. The errors are on the order of the line width.

* martin.marenz@itp.uni-leipzig.de
† wolfhard.janke@itp.uni-leipzig.de

[1] W. Janke (ed.), *Rugged Free Energy Landscapes: Common Computational Approaches to Spin Glasses, Structural Glasses and Biological Macromolecules*, Lect. Notes Phys. **736** (Springer, Berlin, 2008).
[2] K. Hukushima and K. Nemoto, J. Phys. Soc. Japan **65**, 1604 (1996).
[3] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. B **267**, 249 (1991); Phys. Rev. Lett. **68**, 9 (1992).
[4] J. Zierenberg, M. Marenz, and W. Janke, Comput. Phys. Commun. **184**, 1155 (2013).
[5] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1998); **63**, 1195 (1989); **63**, 1658(E) (1989); S. Kumar, J. Rosenberg, D. Bouzida, R. H. Swendsen, and P. Kollman, J. Comput. Chem. **13**, 1011 (1992).