Monte Carlo simulation of crystallization in single-chain square-well homopolymers

Thomas J. Wicks, Jonathan A.D. Wattis and Richard S. Graham[1]
School of Mathematical Sciences, University of Nottingham, Nottingham NG7 2RD, UK.
Dated: July 27, 2020

Abstract: We present Monte Carlo (MC) simulations of the crystallisation transition of single-chain square-well homopolymers, with a continuous description of monomer positions. For long chains with short-ranged interactions this system shows a strong configurational bottleneck, which makes it difficult to explore the whole configuration space. To surmount this problem we combine parallel tempering with a non-standard choice of tempering levels, a bespoke biasing strategy and a method to map results between different temperatures. We verify that our simulations mix well when simulating chains of 128 and 256 beads. Our simulation approach resolves issues with reproducibility of MC simulations reported in prior work, particularly for the transition region between the expanded coil and crystalline region. We obtain highly reproducible results for both the free energy landscape and the inverse temperature, with low statistical noise. We outline a method to extract the free energy barrier, at any temperature, for any choice of order parameter, illustrating this technique by computing the free energy landscape as a function of the Steinhardt-Nelson order parameter for a range of temperatures.
Keywords: Monte-Carlo simulation, single chain, square-well potential

1 Introduction

The collapse of a single polymer chain into a crystal state provides a fundamental model problem for polymer crystallization and protein folding. The crystallisation is controlled by the monomer connectivity and interactions. In computer simulations several classes of interactions are commonly used: lattice models[1–3] and off-lattice interactions using the Lennard-Jones[4–9] or square well potential[10–14]. Bonded interactions can be chosen to give freely jointed[10–12] or semi flexible chains[4,6,8,13–15] Long chains are of particular interest as they approach the chains lengths in realistic systems. However simulating long chains presents computational challenges. Long chains, of 256 monomers, have long been possible for lattice simulations[1–3] and more recently the free energy landscape of the coil to crystal transition for lattice chains of up to 1024 monomers was simulated by dynamic Monte Carlo methods, combined with parallel tempering[16,17]. There has also been recent progress with off-lattice simulations of Lennard Jones chains, where chains of up to 724 monomers have been successfully simulated[5]. However, square well chains with very short ranged interactions are of particular interest because these exhibit a direct coil-crystal freezing transition, whereas

[1]Richard.Graham@nottingham.ac.uk

wider square wells pass first through a collapsed globule phase[11]. These short ranged interactions are the most computationally challenging and have been investigated for 128[11] and 256 monomer chains[18].

The square-well chain model[10] aims to capture features of an isolated homopolymer chain with excluded volume and short range attractive interactions. Typically, segments of the chain are modelled by hard sphere chains that mutually interact via a short-range constant attraction, with no interaction outside this range (see figure 1 and equation (1)). Bonds between monomers are usually of fixed length and freely jointed, or some approximation to this. This system is particularly interesting because it is one of the simplest polymer models that shows crystallisation behaviour[11]. The model has also been linked to protein folding[18]. Thus physical and computational insight into this model system is likely to prove useful to more realistic models of protein folding and polymer crystallisation.

Monte Carlo (MC) simulations of the freely-jointed square well polymer chain have been carried out by Taylor and collaborators[11,19–22]. They used Wang-Landau sampling[23] to comprehensively investigate the temperature-interaction range (T-λ) phase diagram[11]. These authors found three categories of conformational state of the single chain: expanded coil, frozen crystallite and a dense, but amorphous state called the collapsed globule. With decreasing temperature the chain transitions from expanded coil, to collapsed globule and to frozen crystallite. However, for a sufficiently short interaction length ($\lambda < 1.06$), the chain collapses directly from the expanded coil state to the frozen crystallite. Růžička *et al.*[24,25] investigated single square well chains in which the bond length can flex over a very small range. They studied the phase transition dynamics using collision dynamics and forward flux sampling, and compared results to their Wang-Landau simulations of the same system. They extracted a crystallisation temperature from their dynamic simulations by locating the temperature at which the rate of crystallisation matches the rate of melting and showed that temperatures obtained by this method agree with Monte-Carlo simulations[25]. Leitold *et al.*[26,27] also investigated the dynamics of single square well chains. They used a continuous approximation to the true square well potential and sampled the crystallisation transition using molecular dynamics, combined with transition path sampling. They suggested, from analysis of their crystallisation trajectories, a reaction co-ordinate for the crystallisation transition.

This system is problematic even with cutting-edge MC sampling techniques. Taylor *et al.*[11,19–21] showed that, even in Wang-Landau MC simulations, the simulation rarely passes between high and low energy states. Instead, they had to assemble the overall behaviour of the system by joining together separate simulations of the two regions. Consequently, they reported that their results from independent runs are reproducible only outside the transition region, that is at energies intermediate between the coil and crystal states. Similar results were seen in other studies[26]. Furthermore, Růžička *et al.*[25] showed that their static and dynamic results only agree for a particular choice of how to average results over independent runs. Other plausible methods of averaging introduce artefacts that lead to small but noticeable errors in the crystallisation temperature. They also saw somewhat wide variation between independent runs, particularly in the transition region. This variation explains the sensitivity to the averaging method and is likely to be related to the issue of crossing from coil to crystal, as report by Taylor *et al.*[11,19–21]. Finally, Leitold *et al.*[26,27] showed that the choice of reaction co-ordinate to characterise the phase transition dynamics

from coil to crystal, requires careful and detailed analysis of dynamic simulations, leading to a reaction co-ordinate that is a combination of physical quantities that measure the progress towards crystallisation.

The square-well model will be an important system for future studies as it is the simplest macromolecule system that captures the key features of protein folding and polymer crystallisation. Potential future modifications include extension to longer chains and the variation of interaction strength along the chain to correspond to base pairs in proteins. These modifications will probably inherit the problematic simulation features above. Furthermore, the transition region, namely the rare configurations partway between coiled and crystalline configurations, is likely to contain states that are critical to the dynamics of crystallisation and folding. For these reasons there is a strong need for a reliable and efficient way to sample, via MC, the equilibrium distribution of this square well chain system. This MC sampling method should:

1. Cross frequently between high and low energy states from within a single run to produce good sampling of all states and to reduce the variability between independent runs.

2. Have an unambiguous way to average data from multiple independent runs.

3. Retain the ability of Wang-Landau sampling to deliver results at any temperature without running new simulations.

4. Be able to compute the importance of different measures of progress towards crystallisation (and combinations of these) without running new simulations.

In this article, we show that an algorithm combining MC simulation, carefully chosen biasing functions and parallel tempering, when combined with a method to extract the density of states, delivers all of the above points. In particular we employ, a bespoke biasing and tempering strategy to explore and transverse the configurational bottleneck that occurs for long off-lattice chains with short-ranged interactions. The computer code for our algorithm is publicly available[28].

2 Model and Monte Carlo moves

In this section we summarise the interaction model and MC moves used in previous simulations of square well chains[11,24,26].

2.1 Square well chain potential

Non-consecutive particles along the chain i and j, interact through the square well potential,

$$U_{\mathrm{SW}}(r_{ij}) = \begin{cases} \infty & 0 < r_{ij} < \sigma, \\ -\varepsilon & \sigma \leq r_{ij} \leq \lambda\sigma, \\ 0 & r_{ij} > \lambda\sigma, \end{cases} \tag{1}$$

where r_{ij} is the distance between particles i and j, and σ, ε and λ and characterise the hard sphere size, well depth and well width, respectively. The bond angles between neighbouring

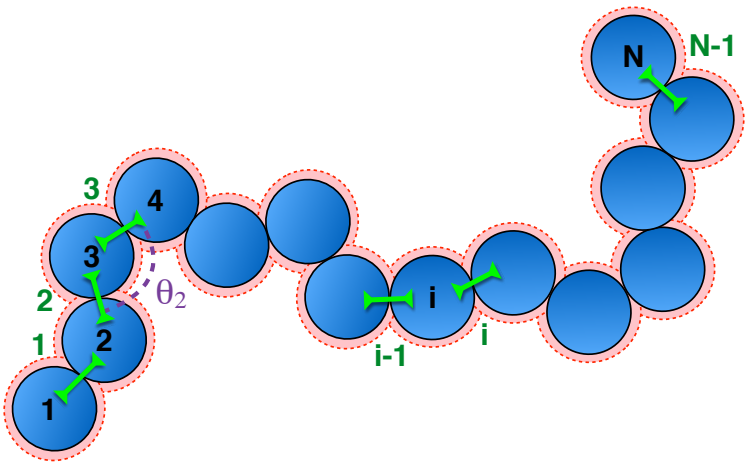


Figure 1: A schematic of the square well chain. Bonds are fixed-length and freely jointed and the particles are hard spheres with a very short-range attractive interaction. Particles are labelled sequentially along the chain, from 1 to N and bond i joins particles i and $i+1$.

particles are freely-rotating (interaction free) and have a fixed length l which is set to σ. Hence the total energy of the chain is a sum over all $U_{\mathrm{SW}}(r_{ij})$ for all non-consecutive particle pairs, and can be written as

$$U_{\mathrm{total}}(\mathbf{r}) = -\varepsilon n_{\mathrm{SW}}(\mathbf{r}), \tag{2}$$

where $n_{\mathrm{SW}}(\mathbf{r})$ is the total number of square-well interactions in configuration $\mathbf{r}$, where $\mathbf{r}$ represents the configuration of all beads in the chain. We nondimensionalise as follows, where * denotes a nondimensional quantity:

$$T = \frac{\varepsilon}{k_{\mathrm{B}}} T^*, \quad U = \varepsilon U^*, \quad r_{ij} = \sigma r_{ij}^*, \tag{3}$$

where k_{B} is Boltzmann's constant.

2.2 Monte Carlo moves

Monte Carlo moves for polymer chains must account for the connectivity and fixed bond length of the neighbouring particles and the excluded volume for all particle pairs. A series of MC moves suitable for square well chains was used by Taylor *et al.*[11] and subsequently adopted by other studies[24,26]. We describe these moves below and adopt them for our MC simulations.

We employ five types of MC move, namely bond pivot, reptation, crank-shaft, end rotation and intramolecular end bridging. In a *bond pivot* move, a bond i is selected at random and a random rotation angle, $\delta\varphi$, chosen uniformly on $[-\delta\varphi_i^{\max}, \delta\varphi_i^{\max}]$, is added to the torsional angle. Bond pivot moves preserve the bond angles, θ_i. Bonds 1 and $N-1$ are discounted from the random draw, as rotating either of these bonds does not change the chain configuration. *Reptation* moves translate the chain along its own length. In a forward reptation move, we detach particle 1 and join it to the opposite end, namely to particle N.

We randomly place particle 1 uniformly on a sphere with radius l, centred on particle N. Similarly, a reverse reptation move is proposed by taking particle N and reattaching it to particle 1, as above. If a reptation move is accepted we relabel all particles so that the particle number runs sequentially along the chain. We perform a *crank-shaft move* by choosing a particle at random, discounting particles 1 and N, and rotating it by $\delta\phi_{\text{crank}}$ about the vector between particles $i-1$ and $i+1$. The rotation angle, $\delta\phi_{\text{crank}}$, is selected randomly from the interval $[-\delta\phi^{\text{max}}_{\text{crank}}(i), \delta\phi^{\text{max}}_{\text{crank}}(i)]$. For *end rotation* moves we randomly select an end particle, which is then rotated about a random axis, passing through the centre of the penultimate bead, by a random angle, $\delta\phi_{\text{end}}$. The random angle, $\delta\phi_{\text{end}}$ is chosen randomly and uniformly from the interval $[-\delta\phi^{\text{max}}_{\text{end}}, \delta\phi^{\text{max}}_{\text{end}}]$. For *end bridging* moves, we summarise the detailed description of end bridging moves by Taylor et al.[11] and Leitold and Dellago[26]. First, we randomly select an end particle, identify all internal particles ($j \in [3, N-2]$) that are within $2l$ of this end and choose one of these at random, denoting this i. The bridging partner, particle i, is then reconnected to the end via removal and reinsertion of the next connected neighbour in the direction of the chosen end, namely particle $i-1$ or $i+1$. The reinserted particle is placed with a randomly chosen azimuthal angle. This process is illustrated in Figure 2. Thus end bridging moves reconnect the chain, changing which particle is the end, while displacing a different single particle. The move is particularly useful for exploring highly crystal states, since they allow small modifications of the crystalline core[26]. Furthermore, a sequence of successful end bridge moves will reposition the chain ends, possibly to a configuration where reptation moves are more likely to be accepted. To ensure that the end bridging moves obey detailed balance, we must accept the moves with probability[11],

$$P_{\text{ac}}(m \to n) = \min\left[1, \frac{b_m R_n}{b_n R_m} \exp(-\Delta U_{\text{total}}/k_B T)\right], \tag{4}$$

where b_m (b_n) denotes the number of possible bridging partners present in state m (n), R_m is the distance of particle i to the previously selected chain end, R_n is the distance of particle i to the new chain end after the move, and ΔU_{total} is the change in the potential energy (2) due to the move. Thus reptation and end bridging moves are complementary. Reptation moves are very effective at guaranteeing complete relaxation of a configuration. However, their acceptance is low when the chain ends are trapped in the crystal interior. End-bridging offers an alternative escape route by enabling the chain ends to diffuse towards the crystal exterior, with minimal re-organisation of the crystal. Acceptance of end-bridging moves is higher for dense configurations. Once at least one chain end is near the crystal exterior then acceptance of reptation moves will improve.

3 Our Monte Carlo simulation algorithm

We simulate a single square well chain, with fixed length, freely rotating bonds using the non-bonded potential and Monte Carlo moves detailed in section 2. Beyond this, our simulation strategy departs from that of Taylor *et al.* and others[24,26]. Rather than Wang-Landau sampling, we perform direct Monte Carlo simulations at the temperature of interest. We use a specific biasing strategy, detailed below, combined with parallel tempering to surmount

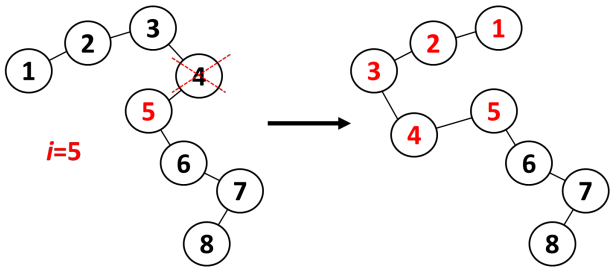


Figure 2: Illustration of the end bridge move. Particle 5 has been selected as the bridging partner for end, 1. Particle 5 is re-bridged to end 1 via removal and reinsertion of particle 4. Finally, particles 1 to 3 are relabelled accordingly.

the bottleneck issues and produce, in a single energy window, good sampling and high reproducibility across the whole space.

3.1 Biasing

We use biasing[29] to obtain more even sampling across all energy states. We choose to bias with respect to the total number of square well neighbours, n_{SW}, which is equivalent to the total energy of the system. We introduce a biasing function $U_{\mathrm{Bias}}(n_{\mathrm{SW}})$ and sample from the biased distribution,

$$P_B \propto \exp\left[-\frac{1}{k_B T}(U_{\mathrm{total}}(\mathbf{r}) - U_{\mathrm{Bias}}(n_{\mathrm{SW}}))\right], \tag{5}$$

where $U_{\mathrm{total}}(\mathbf{r})$ is defined in eqn (2). Thus moves from state m to n are accepted with probability

$$P_{\mathrm{ac}}^{w}(m \to n) = \min\left(1, \Gamma \exp\left[-\frac{1}{k_B T}(\Delta U_{\mathrm{total}}^{m\to n} - \Delta U_{\mathrm{Bias}}^{m\to n})\right]\right). \tag{6}$$

where Γ ensures the correct coefficients for end-bridging moves,

$$\Gamma = \begin{cases} \dfrac{b_m R_n}{b_n R_m}, & m \to n \text{ is an end bridge move,} \\ 1, & \text{otherwise,} \end{cases} \tag{7}$$

as given in equation (4). The results for the occupancy of each energy state n_{SW} sampled from a biased simulation, namely $= P_{\mathrm{Bias}(n_{\mathrm{SW}})}$, can be corrected to the unbiased result[29] by applying

$$P_o(n_{\mathrm{SW}}) = P_{\mathrm{Bias}(n_{\mathrm{SW}})} \exp\left(-\frac{U_{\mathrm{Bias}}(n_{\mathrm{SW}})}{k_B T}\right), \tag{8}$$

where $P_o(n_{\mathrm{SW}})$ and $P_{\mathrm{Bias}(n_{\mathrm{SW}})}$ denote the probability of occupancy of state n_{SW} for the unbiased and biased distributions respectively. Below we detail a method to choose the biasing function to obtain good sampling across the full range of energy states.

3.2 Moves, MC blocks and move acceptances

We divide the simulation into blocks of 10^7 Monte-Carlo steps. At each Monte Carlo step, we propose a single move, choosing randomly from crank-shaft, bond pivot, reptation, end rotation and end bridging moves with probabilities of 0.5, 0.1, 0.1, 0.1 and 0.2, respectively. Every 10^4 moves we vary the maximum angle size for the bond pivot, crank shaft and end rotation moves ($\delta\varphi_i^{\max}$, $\delta\phi_{\rm crank}^{\max}(i)$ and $\delta\phi_{\rm end}^{\max}$, respectively) to target an acceptance ratio of 50% for these moves. The values of $\delta\varphi_i^{\max}$ and $\delta\phi_{\rm crank}^{\max}(i)$ are adjusted independently for each bond i, since bond rotations near the centre of the chain are more likely violate the excluded volume constraints than those near the chain ends, particularly for bond pivot moves.

Reptation moves provide a useful diagnostic measure to monitor effective mixing in our simulation, via *end crossings*. The *reptation number*, $n_{\rm rept}$, records the net number of reptation moves for a particular chain. We iterate $n_{\rm rept}$ by ± 1 whenever a forward/reverse reptation move is accepted. Each time $n_{\rm rept}$ changes by $\pm N$ then the chain has diffused its entire length by reptation, so has fully relaxed. We call these events end crossings. End crossing can only occur when $n_{\rm rept} = kN$, where k is an integer. If the last end crossing occurred at $n_{\rm rept} = kN$ then the next end crossing occurs when $n_{\rm rept} = (k+1)N$ or $n_{\rm rept} = (k-1)N$. Thus, to locate end crossings, we initialise an integer variable, $k^* = 0$, and whenever $n_{\rm rept} = kN$, where $k = k^* + 1$ or $k^* - 1$ we update $k^* = k$ and record that an end crossing has occurred. An end crossing indicates that the chain has renewed its configuration and relaxed any trapped or tightly bound configurations.

3.3 An initial run

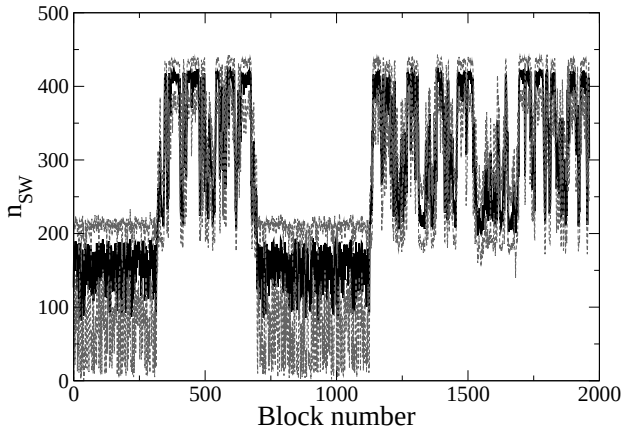


Figure 3: A typical $n_{\rm SW}$ trace plot for a single temperature run at $\overline{T}^* = 0.446$ with biasing aimed at uniform sampling across all $n_{\rm SW}$ states. The thick solid line shows the mean number of square well neighbours, $n_{\rm SW}$, across a single block (10^7 MC moves) and the dashed lines are largest and smallest values of $n_{\rm SW}$ that have at least 10 visits during the block.

In this section we focus on a single temperature simulation and show why this is inade-

quate to explore quantitatively the full space of n_{SW} for long chains. We focus on exploring the energy landscape for the direct transition from expanded coil to crystal at the melting temperature, using parameters identified by Taylor *et al.*[11]. Thus we take $N = 128$, $\overline{T}^* = 0.446$ and $\lambda = 1.05$. Our initial choice for the biasing function attempts to obtain completely uniform sampling across all energy states, by choosing the biasing function to be the estimated free energy of state n_{SW},

$$U_{\mathrm{Bias}}(n_{\mathrm{SW}}) = -k_B \overline{T} \ln(\overline{P}_0(n_{\mathrm{SW}})), \tag{9}$$

where $\overline{T}$ is a reference temperature and $\overline{P}_0(n_{\mathrm{SW}})$ is the unbiased occupancy at $\overline{T}$. For single temperature simulations we take $\overline{T}$ to be the simulation temperature. However, distinguishing between T and $\overline{T}$ becomes important when we introduce parallel tempering. We obtain an initial guess for $\overline{P}_0(n_{\mathrm{SW}})$ either by using literature estimates[19], mapping from our results at a different temperature or by iteratively updating and improving our estimate through a series of simulations.

A typical trace plot for n_{SW} is shown in figure 3, which shows the mean n_{SW} and maximum and minimum values achieved in each block of Monte Carlo steps. This plot shows that initially the simulation explores well the region $0 \leq n_{\mathrm{SW}} \leq 220$, without visiting higher values. This region corresponds to the expanded coil state, extending up to the peak in the free energy barrier, as identified Taylor *et al.*[11]. It is clear that the chain is trapped in the expanded coil state for a prolonged sequence of MC moves. Eventually the simulation passes into the crystal region ($220 \leq n_{\mathrm{SW}} \leq 420$), where it then, similarly, becomes trapped for an extended period. Crossings between the two regions remain very infrequent throughout the entire simulation. Thus we see very good reproducibility of the two regions separately, but poor reproducibility for the transition region around $n_{\mathrm{SW}} \approx 220$ and for the relative importance of these two regions. Even extensive refinement of the biasing function from averages over multiple very long runs does not remove this issue of localised trapping. This suggests that the trapping is not caused by poor acceptance for moves that traverse the transition regions, but instead that such moves are not proposed sufficiently frequently. Our observation are consistent with the results of Taylor *et al.*[11,18,19] who needed to simulate these two regions in separate windows and saw that independent runs are not reproducible in the transition region.

In addition, we see evidence of strong local trapping in the crystalline region. At times, end crossings, as defined in section 3.2, become very infrequent. Instead the reptation number becomes trapped inside a limited range, of size $< N$. This suggests that long-lived crystalline structures persist for many, many MC cycles, a feature that cannot be detected from the reptation move acceptance fraction alone, as this remains high. We speculate that this might be the cause of the somewhat wide variability between independent runs seen in the crystal region by Růžička *et al.*[24,25]. This inability to pass between the coil and crystalline regions may be due to knots, as observed in single chain simulations[4–7], which once formed are very slow to unknot by spontaneous moves in a flat histogram landscape. Longer chains are particularly unlikely to escape such knots.

3.4 Improved algorithm

In section 3.3 above we showed that our initial simulation runs are characterised by strong local trapping and infrequent crossings between the coil and crystalline regions of n_{SW}. This leads to poor reproducibility because frequent crossings between the two regions are needed for efficient sampling of the transition region and to relax long lived crystalline configurations. In this section we detail an algorithm that solves these issues by combining parallel tempering with an unusual choice of tempering levels, a refined biasing strategy and by closely monitoring several indicators of effective mixing.

3.4.1 Parallel tempering

We argued above that the simulation issues with the transition region were due to infrequent proposal of moves that traverse the transition region. Parallel tempering, where multiple copies of the simulation are run in parallel at different temperatures, has the potential to escape such local trapping by proposing swaps between different tempering levels. Thus bottlenecks in the phase space can be bypassed through these swaps. As our single temperature simulations fail to propose sufficiently frequently moves that cross the transition region in both directions then we employ tempering levels both above and below the temperature of interest $\overline{T}$. We set the biasing so that the simulation at $\overline{T}$ occupies the coil and crystalline regions roughly equally. Thus tempering levels below $\overline{T}$ will occupy more frequently the crystalline region and so swaps from $\overline{T}$ to lower tempering levels will provide new pathways into the crystalline region. Similarly, swaps from $\overline{T}$ to higher tempering levels will provide pathways into the expanded coil region.

A further advantage of parallel tempering for this system is that reptation moves are readily accepted for high temperature chains, so these chains will frequently relax fully. Thus the fully relaxed configurations from high tempering levels will propagate to lower tempering levels, recrystallising during this process, and hence will provide a continual supply of fresh crystalline structures that can be explored by the lower tempering levels.

To implement parallel tempering[30] we choose a range of temperatures, $T_1, T_2, \ldots, T_{N_{\mathrm{temp}}}$, where N_{temp} is the total number of tempering levels. One of the T_i is the temperature of interest, $\overline{T}$. We run N_{temp} simulations in parallel using OpenMP, with each temperature having the same biasing function U_{Bias}. Every 10^5 MC steps we pause the simulations and propose n_{swaps} swaps of configurations between adjacent temperatures, chosen at random. Swaps are accepted with probability

$$P_{\mathrm{ac}}(T_i \rightleftarrows T_{i+1}) = \min\left[1, \exp\left(\left(\frac{1}{k_B T_i} - \frac{1}{k_B T_{i+1}}\right)([U^i_{\mathrm{SW}} - U^i_{\mathrm{Bias}}] - [U^{i+1}_{\mathrm{SW}} - U^{i+1}_{\mathrm{Bias}}])\right)\right], \tag{10}$$

where U^i_{SW} and U^i_{Bias} are the square well and biasing potentials, respectively, evaluated using n_{SW} for the simulation at temperature T_i. To ensure effective exchange of configurations between tempering levels, we propose a large number of swaps at each cycle, with n_{swaps} chosen in the range $4N_{\mathrm{temp}}7N_{\mathrm{temp}}$. This is to enable replicas to leap several tempering levels in a single swap cycle. Such leaps are helpful to propagate configurations from above and below the bottleneck into $\overline{T}^*_i$. A wider spacing of tempering levels would still allow these leaps this but at a cost of lower acceptance probabilities for swaps. There is value in breaking large

leaps into smaller sub-leaps as, even if the overall large leap is unsuccessful small subsections of this leap are likely to have been accepted. Once n_{swaps} have been proposed, the N_{temp} simulations are resumed in parallel and the whole process is repeated until the desired number of Monte Carlo steps have been executed. Each tempering level has its own set of maximum rotation angles ($\delta\varphi_i^{\max}$, $\delta\phi_{\text{crank}}^{\max}(i)$ and $\delta\phi_{\text{end}}^{\max}$) which we adjust independently. We implement this parallel tempering algorithm in parallel on multiple CPU cores using OpenMP and our code is freely available[28]. We have verified our code and algorithm by performing the following checks: confirming that the simulation results agree with known analytic results for short chains; confirming that the result for chains of lengths $N = 10, 40, 70$ and 128 are independent of the number and distribution of temperatures and of the choice of biasing function, $U_{\text{Bias}}(n_{\text{SW}})$; and performing simulations for a range of values of $\overline{T}$ and confirming that the results for any tempering level taken from any $\overline{T}$ simulation can be accurately and consistently mapped onto any other temperature using the temperature mapping technique in section 3.6.2.

3.4.2 Biasing choice

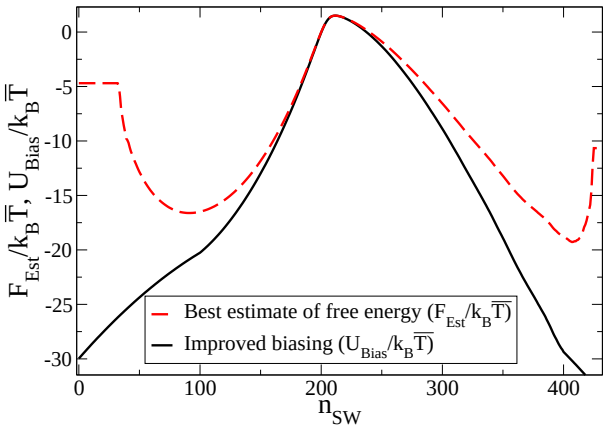


Figure 4: A comparison of the best estimate for the free energy landscape F_{Est} and our biasing approach, U_{Bias}, that leads to improved mixing. The improved biasing holds F_{Est} constant for $n_{\text{SW}} < 100$ and $n_{\text{SW}} > 400$ and subtracts the term $3 \times 10^{-4}(n_{\text{SW}} - 212)^2$ from this clipped version of F_{Est}.

In Wang-Landau sampling[23] the current best estimate of the free energy landscape is chosen as the biasing function (as in eqn (9)) to produce even sampling of all energy levels. However, a key difficulty with simulating this problem is the infrequent proposal of moves that traverse the bottleneck in the transition region. We suggested above that parallel tempering will assist this and, here, we detail a biasing strategy to further promote the proposal of such moves. The idea is to bias the simulation so that it spends more time in the vicinity of the bottleneck, thus creating more frequent traversal moves. Thus we begin with our best estimate for the free energy, as in equation (9) and subtract from this a quadratic

function in n_{SW}, centred on the bottleneck. This revised biasing function is constructed as follows: we define the estimated free energy as

$$\frac{F_{\mathrm{Est}}(n_{\mathrm{SW}})}{k_B\overline{T}} = \begin{cases} -\ln(\overline{P}_0(n_{\mathrm{SW}}^{\mathrm{min}})) & n_{\mathrm{SW}} < n_{\mathrm{SW}}^{\mathrm{min}}, \\ -\ln(\overline{P}_0(n_{\mathrm{SW}})) & n_{\mathrm{SW}}^{\mathrm{min}} < n_{\mathrm{SW}} < n_{\mathrm{SW}}^{\mathrm{max}}, \\ -\ln(\overline{P}_0(n_{\mathrm{SW}}^{\mathrm{max}})) & n_{\mathrm{SW}} > n_{\mathrm{SW}}^{\mathrm{max}}, \end{cases} \tag{11}$$

which is clipped outside the range $n_{\mathrm{SW}}^{\mathrm{min}} < n_{\mathrm{SW}} < n_{\mathrm{SW}}^{\mathrm{max}}$ in order to reduce, but not prevent, exploration of the extreme edges of the energy space; and we then subtract a quadratic function from this to create the new biasing function

$$U_{\mathrm{Bias}}(n_{\mathrm{SW}}) = F_{\mathrm{Est}} - k_B\overline{T}\kappa^*_{\mathrm{Bias}}(n_{\mathrm{SW}} - n^0_{\mathrm{SW}})^2, \tag{12}$$

where the parameter κ^*_{Bias} controls the strength of the localising quadratic about the estimated bottleneck point n^0_{SW}. This biasing strategy is improved, relative to uniform sampling as the simulation spends more time close to the bottleneck and so crosses it more frequently. This approach is illustrated in figure 4 where we compare the best estimate for the free energy, F_{Est}, with an example of our improved biasing strategy, U_{Bias}. For $N = 128$, we used the following values $\kappa^*_{\mathrm{Bias}}/10^{-4} = 1.5, 2.0, 2.5$ and 3.0, with $n^0_{\mathrm{SW}} = 200, 200, 210$ and 212, respectively. A plot of this improved biasing choice is shown in figure 4.

3.5 Verifying good mixing

For polymer chains, the connectivity of the particles means that the acceptance of moves is reduced due to the increased likelihood of particle overlap. Consequently, the simulation is more likely to become trapped in a dense or crystal state, particularly when sampling at a temperature close to or below the melting temperature of the chain. For example, if the ends of the chain are trapped in the centre of a crystal or dense cluster, then a sequence of reptation moves in the same direction is unlikely to be accepted, and so the particular crystal or dense configuration persists. This gives the illusion of achieving good sampling of the trapped state, but in fact correlations in the simulation persist, meaning that we only obtain statistics in the region of the trapped state and good mixing is not achieved (see for example figure 3). Hence we monitor several measures of mixing, as outlined below. Ultimately, it is not possible to show definitely that a simulation has good mixing. Instead we provide strong evidence of this for our simulations by verifying the absence of the signs of poor mixing and by showing a very high level of reproducibility in the results of independent runs.

3.5.1 Diagnostic indicators of poor mixing

The following are indicators of poor mixing in the simulation. Our primary measure comes from monitoring the **trace plots** of n_{SW} against Monte Carlo step. Trace plots, such as figure 3, where n_{SW} remains localised in a limited range for an extended sequence of steps, indicate poor mixing. In contrast, plots that consistently visit the full range of n_{SW} in a just few MC cycles, as in figure 5b, suggest good mixing. We also monitor the **acceptance**

ratio of all move-types. If any move type has very low acceptance, then this suggests potential trapping, or at least inefficient exploration of the phase space. The most important move-type to monitor is reptation, since, as discussed in section 3.2, an extended sequence of reptation moves in the same direction is enough to fully renew the chain configuration. However, the simulation may still be trapped even if the reptation moves have high acceptance ratios, since the reptation number may be varying but only within a narrow range. Thus we monitor the **frequency of end crossings**, as defined in section 3.2. Frequent end crossings ensure that the reptation moves cause full relaxation of prior chain configurations. The above criteria can be monitored *in situ* for each simulation. We also monitor the reproducibility of the simulation outputs but this can only be checked when a series of identical but independent runs near completion.

3.5.2 Handling or bypassing the indicators of poor mixing

For the basic square well chain simulation algorithm, described in sections 2, 3.1-3.3, it is not possible to satisfy the above measures for a single temperature run close to the crystallisation temperature. Instead the trace plots show significant, long-lived trapping (see figure 3). Furthermore, in the crystalline region, both the reptation acceptance ratio and the frequency of end crossings are low. This leads to poor reproducibility in the simulation outputs. In contrast, increasing the temperature improves significantly the reptation number and end crossing frequency, but now only the region of low n_{SW} is explored. Correcting the biasing to sample high n_{SW} states reintroduces the problem of trapping and ineffective reptation.

The above issues lead us to introduce parallel tempering, as described in section 3.4.1. We choose a biasing function matched to the tempering level $\overline{T}^*$, with values of κ^*_{Bias} and n^0_{SW} chosen so that the simulation at $\overline{T}^*$ preferentially visits the vicinity of the n_{SW} value that is hardest to pass through. Simulations at tempering levels above $\overline{T}^*$ explore disproportionately the low n_{SW}, whereas the tempering levels below $\overline{T}^*$ explore the crystalline region. We choose the range and distribution of the tempering levels so that the highest tempering level (hottest chain) has frequent end crossings and the acceptance ratio for swaps between all neighbouring tempering levels is high $(50 - 80\%)$. We also check that individual chain configurations frequently traverse the full range of tempering levels. These measures ensure that new, fully relaxed configurations are continually introduced through the hottest chain and passed through to the coldest chains via swaps and that tightly bound crystalline configurations can consistently be relaxed by swapping up to hotter chains where they readily melt. We additionally verify this by monitoring the trace plot of the chain at $\overline{T}^*$ to ensure that it explores n_{SW} extensively and frequently. To select an appropriate κ^*_{Bias} we begin with a large value to ensure accurate resolution of the free energy in transition region. We then feedback this free energy into the biasing function and gradually decrease the value of κ^*_{Bias} to resolve a wider range of n_{SW} and repeat until κ^*_{Bias} becomes sufficiently small that the reproducibility between independent runs is no longer acceptable. Through this method we find a range of values of κ^*_{Bias} over which a sufficiently wide range of n_{SW} is explored, with reproducible results for the free energy-landscape and all other outputs presented herein, throughout.

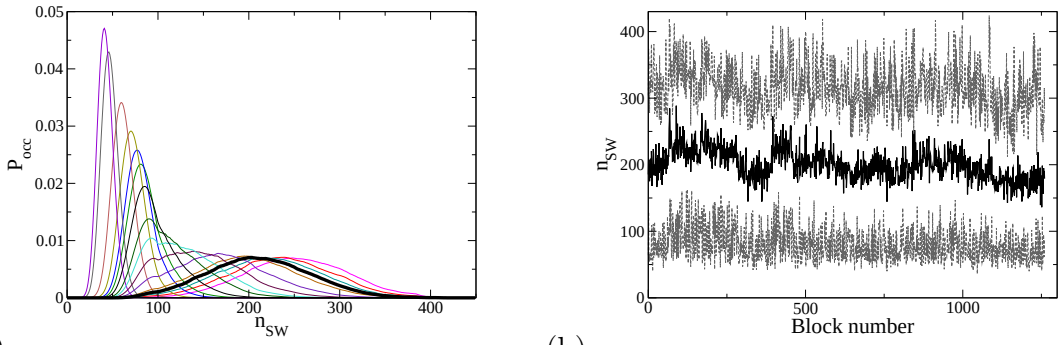


Figure 5: Typical output for the $N = 128$ square-well chain, where $\lambda = 1.05$, $\overline{T}^* = 0.446$, $\kappa^*_{\text{Bias}} = 1.5 \times 10^{-4}$, $N_{\text{temp}} = 16$ and the tempering levels are shown in table 1. (a) Biased occupancies at different tempering levels, with the thicker line being $\overline{T}^*$. (b) Trace plot for $\overline{T}^* = 0.446$. The thick solid line shows the mean n_{SW} value across a single block (10^7 MC moves) and the dashed lines are largest and smallest values of n_{SW} that have at least 10 visits during the block.

3.5.3 Choice of temperatures for parallel tempering

We chose tempering levels as described in section 3.5.2 for the $N = 128$ square-well chain, where $\lambda = 1.05$ and $\overline{T}^* = 0.446$ for $N_{\text{temp}} = 12$ and 16 and the resulting tempering levels are shown in table 1. As the choice of biasing function is designed to produce a broad distribution of n_{SW} at $\overline{T}^*$, then the tempering levels around $\overline{T}^*$ must be closely bunched to keep the swap acceptance high. As the chains become hotter their distribution $P_{occ}(n_{\text{SW}})$ becomes more narrowly peaked in the expanded coil region, meaning that wider gaps in the tempering levels still permit high swap acceptance. Figure 5a shows the resulting biased occupancies for $N_{\text{temp}} = 16$. In figure 5b we see from the trace plot that the simulation at $\overline{T}^*$ visits almost the entire range of n_{SW} in every single MC block.

The general principles guiding our choice of tempering parameters are as follows. We chose the tempering levels and n_{swaps} to ensure frequent propagation of fresh configurations from high and low temperatures towards the temperature of interest, $\bar{T}^*$. In particular, we chose the highest tempering level to ensure sufficiently frequent end crossings (see section 3.2), making the lowest choice that meets this criterion. Similarly, we chose the lowest tempering level to ensure that this replica explored extensively the free energy minimum corresponding to the crystalline state. The intermediate tempering levels were chosen to provide a high swap acceptance ratio (in the range 0.30.9), with higher acceptance ratio chosen around $\bar{T}^*$ to help the simulation explore the bottleneck. We chose n_{swaps} to be in the range $4N_{\text{temp}}7N_{\text{temp}}$. From these highest and lowest temperatures and the desired swap acceptance, this determines the required number of temperatures. We adjusted N_{temp} and the choice of tempering levels, by hand, to meet these criteria. Based on these criteria it would be possible to automate the choice of N_{temp}, n_{swaps} and the tempering levels, potentially utilising the systematic work of Kofke[31] on choosing tempering levels. More specifically, we could combine our current estimate of the density of states (see section 3.6.2) with Kofkes method

Table 1: Tempering levels used for our simulations, chosen using the method in the main text.

Temperature	$N_{\text{temp}} = 16$	$N_{\text{temp}} = 12$
T_1^*	0.4428	0.44280
T_2^*	0.4438	0.4439
T_3^*	0.4445	0.4449
T_4^*	0.4453	0.446 ($= \overline{T}^*$)
T_5^*	0.446 ($= \overline{T}^*$)	0.4519
T_6^*	0.4478	0.4579
T_7^*	0.4496	0.4640
T_8^*	0.4514	0.4701
T_9^*	0.4532	0.4826
T_{10}^*	0.4568	0.4954
T_{11}^*	0.4605	0.5085
T_{12}^*	0.4641	0.5354
T_{13}^*	0.4716	
T_{14}^*	0.4867	
T_{15}^*	0.5178	
T_{16}^*	0.5344	

to predict the swap acceptance fraction to systematically update the choice of tempering levels. However, we leave this extension to future work.

We note here that, if the simulation begins with all chains in the coil state then a long equilibration period, of 200-300 MC blocks, is required, for the coldest chains to reach the crystalline region. Thus we typically begin our simulations using configurations from a prior well-equilibrated run. Simulations begun this way typically equilibrate very quickly, even if we make changes to κ^*_{Bias} and the distribution of tempering levels.

3.6 Analysis of results

For the simulations outlined above, each run takes $\sim$ 5 days on N_{temp} processor cores of 2.6GHz, giving production data of 1300 blocks of 10^7 Monte-Carlo steps. The output data comprises, for each tempering level, a set of occupancy counts for each n_{SW} state, sampled from the biased distribution. In this section we detail how to convert these data into useful outputs.

3.6.1 Averaging over runs

It is useful to average results over several independent runs to decrease statistical noise and to reduce the real-time wait to obtain results. However, Vorselaars et al.[25] showed that, for Wang-Landau sampling of this system, the results are sensitive to the choice of which physical quantity to average over from the independent runs. It is not clear which quantity should be the preferred choice for averaging. The issue arises because each independent Wang-Landau run arrives at a different value for the biasing function. It is not clear how to unambiguously

average over runs with difference biasing. In contrast, our approach allows us to run a series of independent simulations with identical biasing. Thus we can unambiguously average over these runs merely by summing over the number of visits to each n_{SW} state to obtain a more accurate biased distribution, for subsequent analysis. We do this summation separately for each tempering level.

3.6.2 Density of states and mapping between temperatures

Wang-Landau sampling has the desirable feature that results from a single simulation can be mapped to any desired temperature. In this section we use the same underlying physics to perform this temperature mapping for our simulations. As noted in equation (2) the number of crystal neighbours, n_{SW} is directly proportional to the total energy of the chain, so we write

$$P(n_{\mathrm{SW}}, T_i) = \frac{g(n_{\mathrm{SW}})}{Z(T_i)} \exp\left(\frac{\varepsilon n_{\mathrm{SW}}}{k_B T_i}\right), \tag{13}$$

where $g(n_{\mathrm{SW}})$ is the density of states, $Z(T_i)$ is the partition function at temperature T_i and we have used $U_{SW}(n_{\mathrm{SW}}) = -\varepsilon n_{\mathrm{SW}}$. Thus we can extract the logarithm of the density of states from the occupancy of any of the replicas from a parallel tempering run, via

$$\ln g(n_{\mathrm{SW}}) = \ln P(n_{\mathrm{SW}}, T_i) - \frac{\varepsilon n_{\mathrm{SW}}}{k_B T_i} + \ln Z(T_i). \tag{14}$$

Note that $\ln g(n_{\mathrm{SW}})$ is independent of temperature and we determine it only to within an unknown additive constant $\ln Z(T_i)$. From a knowledge of $\ln g(n_{\mathrm{SW}}))$, determined from a simulation at one temperature, T_i, we can obtain predictions at a new temperature $\overline{T}$. For example, for the free energy landscape at $\overline{T}$

$$F(\overline{T}) = -\varepsilon n_{\mathrm{SW}} - k_B \overline{T} \ln g(n_{\mathrm{SW}}), \tag{15}$$

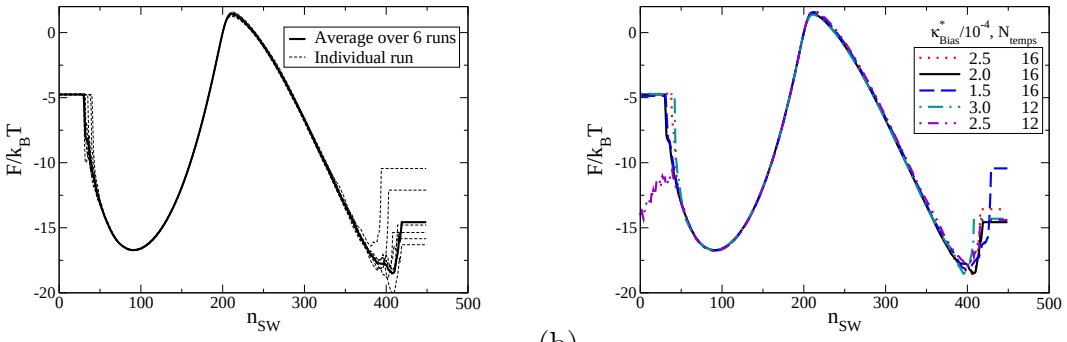
can be calculated from $\ln g(n_{\mathrm{SW}})$ obtained via a simulation at T_i.

4 Results

Having established our simulation algorithm and analysis methods in section 3, in this section we present results. We consider the square-well chain of length $N = 128$ and $\lambda = 1.05$ and we target the crystallisation temperature reported by Taylor et al.[11], $\overline{T}^* = 0.446$. We ran a sequence of simulation runs for this system, using the temperature distributions in table 1 and values of κ^*_{Bias} in the range $1.5 \times 10^{-4} < \kappa^*_{\mathrm{Bias}} < 3.0 \times 10^{-4}$. For each pair of N_{temp} and κ^*_{Bias} we performed six independent, but otherwise identical, repeat runs, to check reproducibility and to average over.

4.1 Results from $\overline{T}$ only

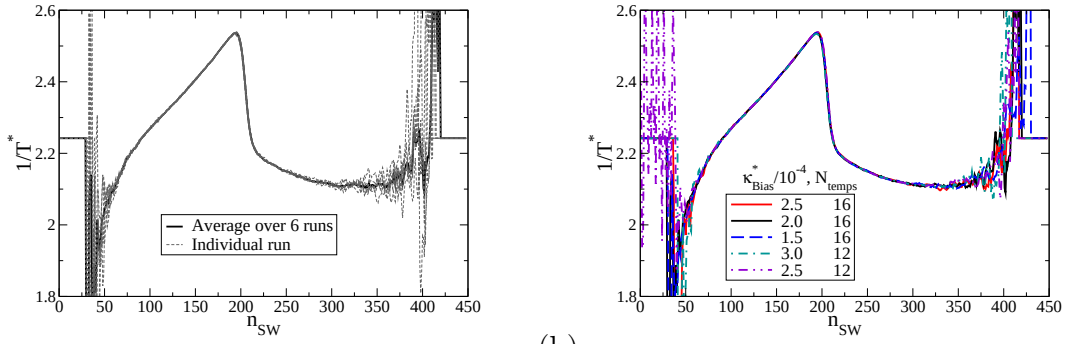
Figures 6 and 7 compare different simulations with results extracted only from the tempering level at $\overline{T}^*$. Figure 6a shows the strong reproducibility of the free energy landscape between independent runs with the same simulation parameters (N_{temp} and κ^*_{Bias}). Figure 6b, where



(a) (b)

Figure 6: Plots of the free energy barrier, illustrating reproducibility for the $N = 128$ square-well chain, where $\lambda = 1.05$ and $\overline{T}^* = 0.446$, using data only from the tempering level $\overline{T}^*$. (a) Six independent runs and their average, for $\kappa^*_{\text{Bias}} = 2 \times 10^{-4}$ and $N_{\text{temp}} = 16$: (see section 3.6.1 for averaging and table 1 for tempering levels). (b) Comparison of results for various values of κ^*_{Bias} and N_{temp} (results are averaged over 6 independent runs).

results are averaged over six independent runs, also shows strong reproducibility in the results from different choices of biasing function and number of tempering levels. Both approaches show particularly strong reproducibility in the low n_{SW} minimum and the transition region maximum. Direct averaging of the barrier heights in figure 6b gives a free energy barrier of $\Delta F = 18.3 k_B T$, with a less than $0.35 k_B T$ variation between the largest and smallest barrier heights in these results. The only significant point that lacks reproducibility is around the minimum in the crystalline region ($n_{\text{SW}} \approx 400$), particularly for larger κ^*_{Bias}. Here there is statistical noise, as the occupancy of the biased distributions is small here due to our choice of biasing function. This could be addressed by further reducing κ^*_{Bias} and performing the clipping of the free energy estimate (see section 3.4.2) at larger $n^{\text{max}}_{\text{SW}}$. However, instead we show in section 4.2, below, how this can be rectified without running further simulations.



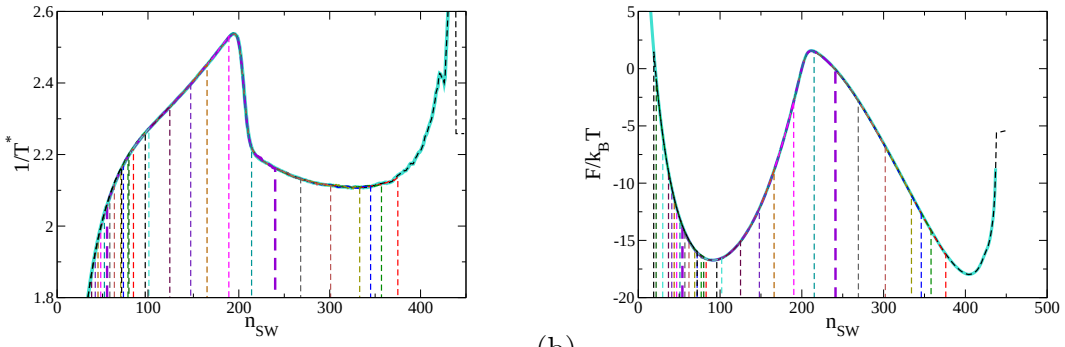
(a) (b)

Figure 7: Reproducibility of the inverse temperature for the $N = 128$ square-well chain, where $\lambda = 1.05$, using data only from the tempering level $\overline{T}^*$. Simulation details and parameters for (a) and (b) are identical to figure 6.

Figure 7 shows results from the same simulations as figure 6 but for the inverse temperature, extracted from the density of states,

$$1/T^* = \frac{\varepsilon}{k_B}\frac{\mathrm{d}S}{\mathrm{d}E} = -\frac{\mathrm{d}\ln g(n_{\mathrm{SW}})}{\mathrm{d}n_{\mathrm{SW}}}, \tag{16}$$

where S is the chain entropy (see reference[11] for further details). The inverse temperature data also show very good reproducibility for most n_{SW} values, especially in the transition region. As in figure 6, there is some statistical noise the around the crystalline minimum ($n_{\mathrm{SW}} \approx 400$) and also at very low n_{SW}.



(a) (b)

Figure 8: Combining results from all tempering levels to cover the full energy range. Results for the $N = 128$ square-well chain, where $\lambda = 1.05$, with $\kappa^*_{\mathrm{Bias}} = 1.5 \times 10^{-4}$ and $N_{\mathrm{temp}} = 16$, with averaging over 6 independent runs. The thin dashed lines are results from individual tempering levels, The thick solid line is the result of stitching the results from the tempering levels together, as described in the text. (a) Inverse temperature and (b) Free energy at $\overline{T}^* = 0.446$.

4.2 Results from combining all tempering levels

By extracting results only from the tempering level at $\overline{T}^*$ we demonstrated very good reproducibility everywhere except for the extreme values of n_{SW}. Here there is statistical noise due to the low occupancy in the biased distribution for the tempering levels at $\overline{T}^*$. However, for other tempering levels, these extreme states are well visited, see for example figure 5(a). We can map data from these other temperatures to $\overline{T}^*$ using the methods in section 3.6.2. From this we obtain well-resolved data for all relevant n_{SW} and also test the internal consistency of our results from different tempering levels. We removed noisy data from each individual tempering level by excluding states n_{SW} for which the biased occupancy for tempering level i was below 0.03% of all MC steps. We retained all data for the low n_{SW} states for the hottest chain and the high n_{SW} states for the coldest chain. Figure 8 shows the results from all 16 tempering levels mapped from T^*_i to $\overline{T}^* = 0.446$ via equation (14). This shows the strong internal consistency between mapped results, demonstrated by the high level of overlap between results from different tempering levels. Individual tempering levels generally provide

data over a wide range of n_{SW} states, typically > 150 states, with even more overlap for levels close to $\overline{T}^*$. To illustrate this, the vertical dashed lines in figure 8 show the edges of the data obtained from each tempering level and one tempering level is highlighted with a thicker dashed line. Thus the high degree of superposition at each n_{SW} state shows, in many cases, very close agreement between > 10 tempering levels.

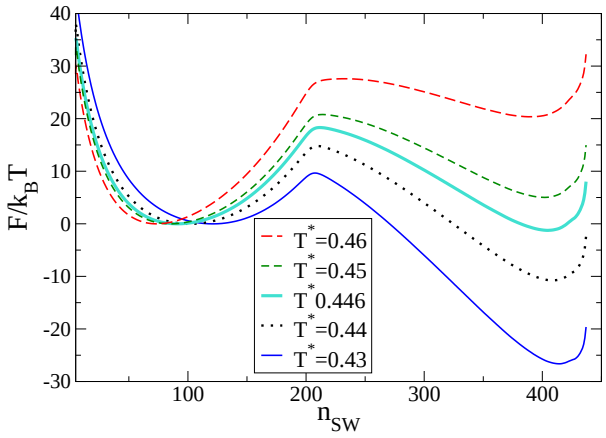
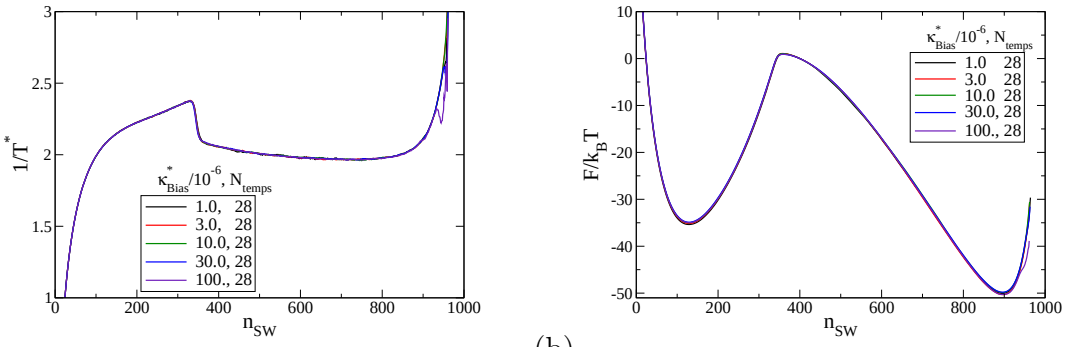


Figure 9: The free energy landscape mastercurve from figure 8 mapped to a range of new temperatures using the technique from section 3.6.2.

From data obtained at each tempering level, we can stitch together a mastercurve that has good statistics at all relevant n_{SW} values. This mastercurve is shown as the thicker solid lines in figure 8. We obtained this curve as follows: we define the 'maximin' point $n_{\mathrm{SW}}^{(i)}$ as the value of n_{SW} at which $\min(P_w(n_{\mathrm{SW}}, T_i), P_w(n_{\mathrm{SW}}, T_{i+1}))$ is maximised; the mastercurve then takes the value obtained from tempering level i for the range $n_{\mathrm{SW}}^{(i-1)} < n_{\mathrm{SW}} < n_{\mathrm{SW}}^{(i)}$. We obtain the free energy values only to within an unknown additive constant so we add a constant to the free energy from each tempering level so that $F_{T_i}(n_{\mathrm{SW}}^{(i)}, \overline{T}) = F_{T_{i+1}}(n_{\mathrm{SW}}^{(i)}, \overline{T})$, where $F_{T_i}(n_{\mathrm{SW}}, \overline{T})$ denotes the free energy state n_{SW} obtained by mapping tempering level T_i to the reference temperature $\overline{T}$. From the master curve for $1/T^*$ in figure 8a we used the equal area construction[11] to compute the phase transition temperature and obtained $T^* = 0.447$, essentially the same value as Taylor *et al.*[11]. Having obtained a mastercurve for the free energy landscape at $\overline{T}^* = 0.446$, we can map this to any new temperature by using the technique in section 3.6.2, as illustrated in figure 9.

4.3 Results for a chain of $N = 256$

Taylor *et al*[18] studied $N = 256$ with $\lambda = 1.05$. They reported significant variation between independent runs, particularly at the barrier peak. We ran simulations at $\overline{T}^* = 0.477$ using our approach above and our simulations reproducibly converged to consistent results for both $1/T^*$ and the free energy barrier, across the whole range of n_{SW} (see figure 10). To obtain reproducible runs we used 28 tempering levels, chosen above and below $\overline{T}^*$. We verified



(a) (b)

Figure 10: Results for the $N = 256$ square-well chain, where $\lambda = 1.05$, averaged over 4 independent runs, obtained by stitching the results from the tempering levels together, as described in the text. (a) Inverse temperature and (b) Free energy at $\overline{T}^* = 0.477$.

that all of the indicators of good mixing in section 3.5 are satisfied, that all values of n_s are well-sampled by multiple tempering levels and that results from these tempering levels are consistent when mapped to the same temperature as in section 4.2. For each value of κ^*_{Bias} we averaged over four independent runs, each of ~ 5 day. However, each individual run is very close to the final average, in a similar manner to figure 6(a).

5 Alternative reaction co-ordinates

Leitold *et al.* used dynamic simulations of square well chains to derive reaction co-ordinates for the coil-crystal transition[26,27]. They showed that the number of square well neighbours, n_{SW}, does not correlate sufficiently strongly to progress towards the transition, and so it is not a suitable reaction co-ordinate. Instead they suggested a combination of n_{SW} together with an adaption of the Steinhardt-Nelson order parameter, q_8, which measures crystalline order by considering both the number and spatial position of near-neighbour particles[32]. Initial studies[26] of this problem led to a reaction coordinate that is a linear combination of n_{SW} and q_8 and later work refined this to a non-linear function[27]. It is clear that future work may further refine this order parameter, possibly by introducing further quantities into the reaction co-ordinate. Thus, there is a need for a flexible method to obtain the free energy landscape as a function of any specified reaction co-ordinate. We develop such a method in this section.

An adaption of the Steinhardt-Nelson order parameter introduced by Leitold and Dellago[26] is as follows. For each pair of particles that are closer than the distance $r_c = 1.05\sigma$, compute the connection coefficients $d_6(i, j)$, via equation (20) in appendix A. Two particles i and j are *connected* if $d_6(i, j) \geq 0.5$. A particle is labelled *crystalline* if it has at least 5 neighbours and its number of non-connected neighbours is no more than 1, and n_C denotes the total number of crystalline particles. This definition identifies particles near the surface of a crystal structure as crystalline, while also avoiding incorrectly defining particles in the core of a dense, unordered region as crystalline.

5.1 Simulating the occupancy of n_X via the density of states

Here we consider an arbitrary order parameter n_X. At temperature $\overline{T}$, the probability of state n_X is given by

$$P(n_X, \overline{T}) = \sum_{i=0}^{N_E} P(E_i, \overline{T}) P(n_X | E_i), \tag{17}$$

where N_E is the total number of available energy states, $P(E_i, \overline{T})$ is the probability energy state E_i at temperature $\overline{T}$ and $P(n_X|E_i)$ is the conditional probability of the system being in state n_X, given that the configuration has energy E_i. $P(E_i, \overline{T})$ is known from the density of states, by eqn (13). Hence

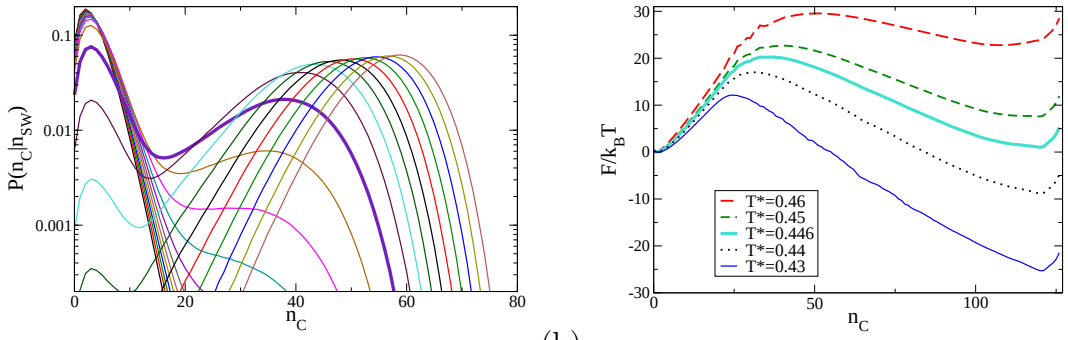
$$P(n_X, \overline{T}) = \frac{1}{Z(\overline{T})} \sum_{n_{\mathrm{SW}}=0}^{N_E} g(n_{\mathrm{SW}}) \exp\left(\frac{\varepsilon n_{\mathrm{SW}}}{k_B \overline{T}}\right) P(n_X | n_{\mathrm{SW}}), \tag{18}$$

where we have used $E_i = -\varepsilon n_{\mathrm{SW}}$. The density of states $g(n_{\mathrm{SW}})$ can be obtained from a simulation at any temperature (or a combination of temperatures as in section 4). We note that $g(n_{\mathrm{SW}})$ can only be extracted to within a multiplicative constant and that, in general, $Z(\overline{T})$ is not known, but these can both be corrected for by renormalising the distribution of $P(n_X, \overline{T})$ obtained from equation (18). Thus the only unknown element in eqn (18) is $P(n_X|n_{\mathrm{SW}})$, which is independent of temperature as it considers configurations at fixed total energy.

We obtain $P(n_X|n_{\mathrm{SW}})$ during a simulation by storing sets of configurations grouped by n_{SW}, which can then be converted to histograms of $P(n_X|n_{\mathrm{SW}})$ during post-processing. By saving a sample of configurations at each energy level, many difference choices of n_X can be performed during post-processing without running further simulations. Since $P(n_X|n_{\mathrm{SW}})$ is independent of T then we can collect and combine data from each tempering level to rapidly obtain good statistics across all energy levels.

5.2 Results for the Steinhardt-Nelson order parameter

We computed the Steinhardt-Nelson order parameter, n_C, from the sequence of runs outlined in section 4, binning the data by energy level n_{SW} so that we could construct histograms of $P(n_C|n_{\mathrm{SW}})$. We confirmed that the resulting $P(n_C|n_{\mathrm{SW}})$ distributions are independent of tempering level and choice of biasing function. Results are shown in figure 11a. For low and high values of n_{SW} the distribution of n_C is typically fairly narrow, indicating a reasonably close correlation between n_{SW} and n_C. However in the transition region ($n_{\mathrm{SW}} \approx 200$) the distribution broadens and becomes bimodal, with peaks corresponding to crystalline and amorphous configurations. Intriguingly, $P(n_C|n_{\mathrm{SW}})$ is low around $n_C \approx 16$ for all values of n_{SW}. Figure 11b shows the free energy landscape for several temperatures, computed via eqn (18). Here $P(n_C|n_{\mathrm{SW}})$ data from figure 11a is used, along with the $g(n_{\mathrm{SW}})$ data used to compile figure 8. This gives a free energy barrier of $\Delta F = 20.25 k_B T$ at $T^* = 0.446$, with a critical nucleus size of $n_C = 37$.



(a) (b)

Figure 11: Results for n_C for the $N = 128$ square-well chain, where $\lambda = 1.05$, with $\kappa^*_{\text{Bias}} = 1.5 \times 10^{-4}$ and $N_{\text{temp}} = 16$. (a) Distributions of $P(n_C|n_{\text{SW}})$ for $n_{\text{SW}} = 150...250$ in increments of 5. The thicker line corresponds to $n_{\text{SW}} = 205$. (b) The free energy landscape for n_C for a range of temperatures, computed via eqn (18).

6 Discussion and Conclusions

We studied the square-well chain model, choosing model parameters that are known to give a direct first order transition from the expanded coil to crystalline state. For long chains, this system show a configuration bottleneck in the transition region between the coil and crystal state. We surmount the this bottleneck through a combination of biasing and parallel tempering. Although biasing and tempering have previously been applied to this problem for lattice chains[16,17], off-lattice chains are considerably more challenging. We produced a specifically designed biasing and tempering strategy to address the bottleneck problem when simulating off-lattice chains. Our simulation approach resolves reproducibility issues reported in previous MC simulations, particularly for the transition region between the expanded coil and crystalline region. We obtained highly reproducible results for both the free energy landscape and the inverse temperature, both expressed as a function of the total number of near neighbour interactions n_{SW}. By extracting the density of states from our simulation at the freezing temperature, we could map our results to any temperature. We described a method to extract the free energy barrier, at any temperature, for any order parameter that can be computed from the instantaneous chain configuration. We illustrated this technique by computing the free energy landscape as a function of the Steinhardt-Nelson order parameter, n_C, for a range of temperatures.

We found that for long square well chains, there are inherent difficulties related to mixing that arise due to the connectivity of particles, which leads to self-entanglement and an excluded volume bottleneck. We tackle these issues using parallel tempering with a non-standard choice of tempering levels, where we simulate systems at tempering levels *lower* than the temperature of interest, $\overline{T}$, and increase the density of tempering levels close to $\overline{T}$. Tempering levels below $\overline{T}$ offer swaps from the amorphous region into the crystalline region and assist in equilibrating the crystalline region. The high density of tempering levels around $\overline{T}$ is required because the biasing is chosen so that $P_{occ}(n_{\text{SW}})$ very broad at $\overline{T}$.

We defined several measures of mixing that we monitored throughout. These include

monitoring the acceptance ratios of configuration swaps between tempering levels, the evolution of the energy traces, and the reptation number, to ensure effective de-correlation of configurations. We further used biasing, choosing a biasing function encourage detailed exploration of the bottleneck in the transition region whilst still adequately visiting other regions of the phase space. Unlike previous Wang-Landau approaches to this problem, our technique allowed us to use the same biasing over several independent runs. This means that multiple runs can be averaged over in an unambiguous way, leading to improved accuracy without introducing artefacts from the averaging method. We extracted from our simulation data the density of states, which showed very extensive overlap agreement between results from different tempering levels. Thus, for each energy state in turn we selected the result from the tempering level with the best sampling to obtain high-quality mastercurve, spanning the full range of energy states. Combining these techniques leads to highly reproducible results with low stochastic noise for chains of 128 and 256 beads, particularly in the very challenging transition region.

There are immediate applications of our algorithms to longer homopolymer chains, heteropolymers, where the interactions along the chain are non-uniform, and stiff chains, having a potential associated with their torsional angles, to give a simple model of proteins. In all cases, the ability of our algorithms to explore effectively the energy space and by pass bottlenecks will be highly useful, as will the methods to map results to any desired temperature. Adding external forces to the chain in our simulations will produce toy-models to aid the understanding of protein folding/unfolding under mechanical forces[33] and flow-induced crystallisation in polymers[34,35]. Simulations of the crystallisation dynamics[24–27] and the reaction co-ordinates identified by Leitold *et al.*[26,27], could be combined with techniques to project high-dimensional barrier crossing simulations onto low-dimensional co-ordinates[36,37], to produce an analytically tractable model for the crystallisation dynamics of this system. Accurate simulations of the free energy landscape, especially at the barrier peak, are a prerequisite for this approach.

Acknowledgements

We thank Dr. Bart Vorselaars for useful discussions concerning this work. We acknowledge funding from the EPSRC, in the form of a PhD studentship (TJW, award number 1091699) and a Design by Science grant (RSG, EP/P005403/1). We acknowledge access to the University of Nottingham High Performance Computing Facility and the use of Athena at HPC Midlands+, which was funded by the EPSRC on grant EP/P020232/1, as part of the HPC Midlands+ consortium.

A Defining connected crystal particles via the Steinhardt-Nelson method

Crystal-like connections between nearby particles can be defined as follows[38]. For each particle i a bond orientational order parameter is given by

$$\overline{q}_{6m}(i) = \frac{1}{N_b(i)} \sum_{j=1}^{N_b(i)} Y_{6m}(\theta_{ij}, \phi_{ij}), \tag{19}$$

where the sum runs over all particles j that are square well neighbours of particle i, Y_{lm} are spherical harmonics and θ_{ij} and ϕ_{ij} are the polar angles of the vector between particles i and j with respect to some reference coordinate system. The following complex dot product is then defined for each neighbouring pair of particles i and j

$$d_6(i,j) = \frac{\sum_{m=-6}^{6} \tilde{q}_{6m}(i)\tilde{q}^*_{6m}(j)}{\left(\sum_{m=-6}^{6} |\tilde{q}_{6m}(i)|^2\right)^{1/2} \left(\sum_{m=-6}^{6} |\tilde{q}_{6m}(j)|^2\right)^{1/2}}, \tag{20}$$

where * denotes complex conjugate. Particles i and j are connected if their $d_6(i,j)$ exceeds a certain threshold d_c, for which we use a value of 0.5.

References

[1] F. Rampf, K. Binder, and W. Paul, Journal of Polymer Science Part B: Polymer Physics **44**, 2542 (2006).

[2] W. Janke and W. Paul, Soft Matter **12**, 642 (2016).

[3] M.-B. Luo, R.-R. Tian, X. Yang, J.-H. Huang, and M. Mamat, Journal of Polymer Science Part B: Polymer Physics **56**, 1053 (2018).

[4] W. Janke and M. Marenz, in *Journal of Physics: Conference Series* (IOP Publishing, 2016), vol. 750, p. 012006.

[5] M. Marenz and W. Janke, Phys. Rev. Lett. **116**, 128301 (2016).

[6] W. Janke, M. Marenz, and J. Zierenberg, Lobachevskii Journal of Mathematics **38**, 978 (2017).

[7] S. Majumder, J. Zierenberg, and W. Janke, Soft Matter **13**, 1276 (2017).

[8] J. Wu, C. Cheng, G. Liu, P. Zhang, and T. Chen, J. Chem. Phys. **148**, 184901 (2018).

[9] K. Qi, B. Liewehr, T. Koci, B. Pattanasiri, M. J. Williams, and M. Bachmann, J. Chem. Phys. **150**, 054904 (2019).

[10] Y. Zhou, C. K. Hall, and M. Karplus, Phys Rev Lett **77**, 2822 (1996).

[11] M. P. Taylor, W. Paul, and K. Binder, J. Chem. Phys. **131**, 114907 (2009).

[12] A. Böker and W. Paul, The European Physical Journal E **39**, 5 (2016).

[13] S. V. Zablotskiy, J. A. Martemyanova, V. A. Ivanov, and W. Paul, Polymer Science Series A **58**, 899 (2016).

[14] T. Škrbić, J. R. Banavar, and A. Giacometti, J. Chem. Phys. **151**, 174901 (2019).

[15] T. Škrbić, T. X. Hoang, A. Maritan, J. R. Banavar, and A. Giacometti, Soft Matter **15**, 5596 (2019).

[16] W. Hu, D. Frenkel, and V. B. F. Mathot, The Journal of Chemical Physics **118**, 3455 (2003).

[17] W. B. Hu and D. Frenkel, The Journal of Physical Chemistry B **110**, 3734 (2006).

[18] M. P. Taylor, W. Paul, and K. Binder, Phys. Rev. E **79**, 050801(R) (2009).

[19] M. P. Taylor, W. Paul, and K. Binder, Physics Procedia **4**, 151 (2010).

[20] M. P. Taylor, W. Paul, and K. Binder, Polymer Science Ser. 3 **55**, 23 (2013).

[21] M. P. Taylor and P. P. Aung, Phys Rev. E **88**, 012604 (2013).

[22] M. P. Taylor, W. Paul, and K. Binder, J. Chem. Phys. **145**, 174903 (2016).

[23] F. Wang and D. P. Landau, Phys Rev. Lett. **86**, 2050 (2001).

[24] Š. Růžička, D. Quigley, and M. P. Allen, Phys. Chem. Chem. Phys. **14**, 6044 (2012).

[25] B. Vorselaars, Š. Růžička, D. Quigley, and M. P. Allen, Phys. Chem. Chem. Phys. **15**, 21101 [arXiv:1611.00386] (2013).

[26] C. Leitold and C. Dellago, J. Chem. Phys. **141**, 134901 (2014).

[27] C. Leitold, W. Lechner, and C. Dellago, J. Phys.: Condens. Matter **27**, 194126 (2015).

[28] T. J. Wicks and R. S. Graham, *A Monte Carlo code for the square well chain model with biased sampling and parallel tempering* (2020), URL `https://github.com/RichGraham/SqWellChain`.

[29] M. Allen and D. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1989).

[30] D. J. Earl and M. W. Deem, Phys. Chem. Chem. Phys. **7**, 3910 (2005).

[31] D. A. Kofke, The Journal of Chemical Physics **117**, 6911 (2002).

[32] P. Steinhardt, D. Nelson, and M. Ronchetti, Phys. Rev. B **28**, 784 (1983).

[33] D. J. Brockwell, E. Paci, R. C. Zinober, G. S. Beddard, P. D. Olmsted, D. A. Smith, R. N. Perham, and S. E. Radford, Nat Struct Biol **10**, 731 (2003).

[34] R. S. Graham, Chem. Commun. **50**, 3531 (2014).

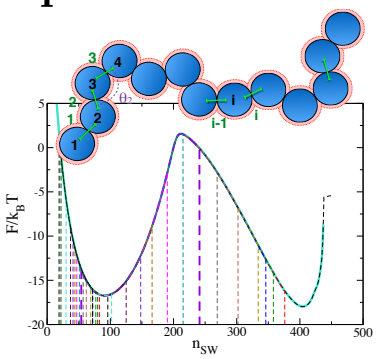
[35] R. S. Graham, J Rheol **63**, 203 (2019).

[36] M. J. Hamer, J. A. D. Wattis, and R. S. Graham, Soft Matter **8**, 11396 (2012).

[37] D. J. Read, C. McIlroy, C. Das, O. G. Harlen, and R. S. Graham, Phys. Rev. Lett. **124**, 147802 (2020).

[38] L. Filion, M. Hermes, R. Ni, and M. Dijkstra, J. Chem. Phys. **133**, 244115 (2010).

Graphical Abstract



We simulate the free energy landscape of a single square-well chain for chains of 128 and 256 monomers. Our Monte-Carlo algorithm leads to a very high level of reproducibility, across the whole landscape. This landscape determines the crystallization temperature and the activation barrier to crystallization.