Population annealing molecular dynamics with adaptive temperature steps

Christiansen, H., Weigel, M. & Janke, W.

Original citation:
Christiansen, H, Weigel, M & Janke, W 2019, 'Population annealing molecular dynamics with adaptive temperature steps' Journal of Physics: Conference Series, vol. 1163, 012074.
https://dx.doi.org/10.1088/1742-6596/1163/1/012074

DOI 10.1088/1742-6596/1163/1/012074
ISSN 1742-6588
ESSN 1742-6596

Publisher: IOP Publishing

Population annealing molecular dynamics with adaptive temperature steps

IOP Conf. Series: Journal of Physics: Conf. Series 1163 (2019) 012074 doi:10.1088/1742-6596/1163/1/012074

Population annealing molecular dynamics with adaptive temperature steps

Henrik Christiansen[1], **Martin Weigel**[2] **and Wolfhard Janke**[1]

[1] Institut für Theoretische Physik, Universität Leipzig, Postfach 100 920, 04009 Leipzig, Germany

[2] Applied Mathematics Research Centre, Coventry University, Coventry CV1 5FB, England

E-mail:
henrik.christiansen@itp.uni-leipzig.de
martin.weigel@complexity-coventry.org
wolfhard.janke@itp.uni-leipzig.de

Abstract. Population annealing is a novel generalized-ensemble simulation scheme used in large-scale parallel Monte Carlo simulations of disordered spin systems and similar problems. In a recent publication we proposed a generalization of this method to molecular dynamics simulations of biopolymers. In the present article we review this work and introduce a scheme for automatically choosing the temperature steps based on the observed distribution of potential energies in the population of simulated replica.

1. Introduction

Due to a rugged free-energy landscape, studies of the folding behavior of peptides are among the computationally most difficult problems in computer simulations. In such problems, the systems get trapped in local minima that are separated by high barriers in free energy that are increasingly difficult to overcome by activated dynamics as the temperature is gradually lowered. Several numerical approaches have been developed to overcome such trapping problems and the ensuing loss of ergodicity in simulations of biopolymers. Parallel tempering (PT), sometimes also known as replica exchange, has previously been shown to have the potential of successfully sampling the full configuration space when combined with molecular dynamics (MD) simulations of peptides [1, 2]. In PT, one considers a number of replicas ($\leq 10^2$) of the same system, which are, a priori, simulated independently at different temperatures, ranging from unphysiologically high temperatures where configurations equilibrate quickly, forgetting their previous history, down to the physiologically relevant temperatures, where the peptide is in a folded state. The simulations running at neighboring temperatures are subsequently permitted to exchange configurations according to an acceptance probability taking into account the Boltzmann weight at the two temperatures. This setup results in a diffusive motion of replicas in temperature space, thus allowing the replicated system to escape from local free-energy minima by heating copies up to the highest simulated temperatures before they cool down again, typically arriving in a different local minimum from the one occupied previously. Although PT can easily be implemented in parallel, the number of processors which can effectively be used is limited, as a setup with too many closely spaced temperatures does not improve the equilibration and even slows down the random walk in temperature space.

In contrast, population annealing (PA) [3, 4, 5, 6] is a simulation scheme introduced in the context of Monte Carlo (MC) simulations that can make use of very large numbers of parallel workers and hence physical cores and massively parallel accelerators such as GPUs [7] without ensuing problems of bad parallel scaling or decreasing simulational efficiency. In this method one considers a multitude (typically at least 10^4) of independent replicas which are initially equilibrated at a high (or potentially even infinite) temperature to get an ensemble of configurations that is perfectly or nearly perfectly uncorrelated. The temperature is then lowered in a sequence of steps, similar to what is done in simulated annealing for a single replica [8]. However, in contrast to simulated annealing, the population is additionally resampled to the new temperature after a number of simulation steps. This resampling leads to a significant speed-up in equilibration and eliminates energetically stuck configurations (reminiscent of the "go with the winners" strategy of Ref. [9]). Observables in PA are calculated using an average over the ensemble, whereas in PT averages are taken over the time series at a fixed temperature. Population annealing has been shown to perform well, e.g., in the simulation of spin glasses [4, 6]. We recently showed how PA can be combined with MD (PAMD) to efficiently simulate macromolecular systems such as the folding of the peptide met-enkephalin [10]. In the present work we demonstrate how the schedule of temperature steps in this scheme does not need to be pre-determined, but it is much preferable to decide about the next temperature to include as one goes along by utilizing the information about the distribution of potential energies resulting from the parallel simulation of a population of replicas.

2. Model and method

The algorithm for PAMD as described in Ref. [10] comprises the following steps:

(i) Set up an equilibrium ensemble of R independent copies of the system at some high temperature T_0.

(ii) Choose the next temperature T_i from a pre-defined sequence.

(iii) Resample the ensemble of systems to the new temperature $T_i < T_{i-1}$ by replicating each copy a number of times proportional to the relative Boltzmann weight $\tau_j \propto e^{-(1/k_B T_i - 1/k_B T_{i-1})E_j}$, where E_j is the potential energy of the jth replica. Similarly to PT, the momenta are easily adjusted to the new temperature by rescaling $p_k \to \sqrt{\frac{T_i}{T_{i-1}}} p_k$.

(iv) Update each copy with θ simulation steps of the underlying MD algorithm.

(v) Calculate observables $\mathcal{O}$ at temperature T_i as population averages.

(vi) Goto step (ii) until T_i reaches or falls below the target temperature T_N.

In our previous work [10], studying the folding of the penta-peptide met-enkephalin with the amino-acid sequence Tyr-Gly-Gly-Phe-Met, we adapted a temperature set ($T_i = 700$, 585, 489, 409, 342, 286, 239, 200 K) from a PT simulation of the same peptide [2]. However, finding an appropriate temperature set in PT simulations is a non-trivial task, for which a number of solutions are being discussed [11, 12, 13]. It is clear, for instance, that too large temperature steps in PT lead to low acceptance rates of swap moves, whereas too many small steps lead to a slowing down resulting from the scaling properties of the temperature random walk. Similar difficulties arise in PAMD with a fixed temperature protocol. In population annealing, however, it is possible to exploit the fact that one has access to the energy histogram at temperature T_{i-1} from the population of replicas at the point of having to decide about the next step, T_i, in the temperature scheme. To benefit from this additional information, we apply single-histogram reweighing [14] to get an estimate of the energy histogram at T_i. We then require that the histograms of potential energy at T_i and T_{i-1} have a certain prescribed overlap $\alpha^\star$ at each step. The actual overlap can be estimated before taking the real temperature step from the following

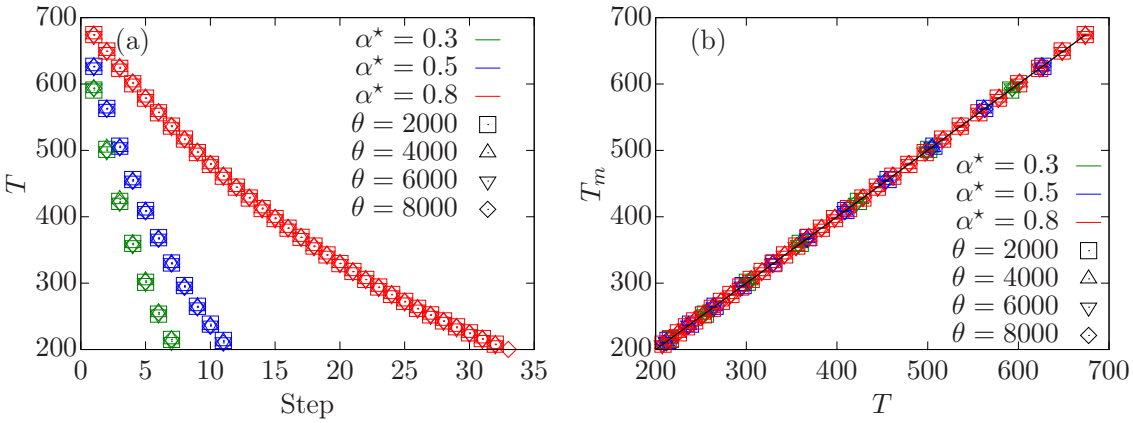

Figure 1. (a) Temperatures found by the PAMD simulation with adaptive temperature stepping are shown for different values of $\alpha^\star$ and number θ of MD steps. The symbols encode the MD steps θ and the colors relate to the values of $\alpha^\star$. (b) The observed temperature T_m (measured just before the resampling step) during the simulations is shown against the heatbath temperature T for the same values of θ and $\alpha^\star$ as in (a).

expression [7],

$$\alpha(T_{i-1}, T_i) = \frac{1}{R}\sum_{j=1}^{R} \min\left(1, \frac{\exp[-(1/k_B T_i - 1/k_B T_{i-1})E_j]}{Q(T_{i-1}, T_i)}\right). \quad (1)$$

Here E_j is the energy of the jth replica and $Q = \sum_{j=1}^{R} \exp[-(1/k_B T_i - 1/k_B T_{i-1})E_j]$ serves as a normalization. Before the simulation one then only has to choose a fixed value of $\alpha = \alpha^\star$, and any given value of $\alpha^\star$ results in a certain temperature sequence. In the above algorithm, we hence propose to replace step (ii) by the following adaptive variant:

(ii') Find the next temperature T_i in such a way, that $\alpha(T_{i-1}, T_i) = \alpha^\star$.

The question of what are reasonable values of $\alpha^\star$ is investigated in the following section.

In the simulations reported below we used OpenMM [15] to perform the MD, relying on the Langevin thermostat and a velocity-Verlet like integrator. The interactions of the peptide are modeled via the AMBER force-field ff94 [16]. It is worthwhile to point out that for PAMD simulations to work properly the thermostat has to be stochastic as otherwise all of the resampled copies of a population member will follow exactly the same trajectory in further evolving the system, thus defying the decorrelating purpose of the resampling step.

3. Results

We fixed the relevant PAMD parameters at $R = 10\,000$ replicas, a time step of $\mathrm{d}t = 0.5$ fs, and a friction coefficient $\gamma = 1/\mathrm{ps}$. For the constant energy overlap simulations, we chose $T_0 = 700$ K as the initial high temperature and set $T_N = 200$ K as the target temperature. The start configurations at T_0 were taken from a canonical simulation of length 25 ns, followed by short simulations for each replica to further decorrelate the population. To find a suitable set of overlaps $\alpha^\star$ and corresponding number θ of MD simulation steps per temperature, we tested $\alpha^\star = 0.3, 0.5, 0.8$ and $\theta = 2000, 4000, 6000, 8000$. In Fig. 1(a) we show the temperatures realized by the constant energy overlap PAMD runs by choosing $\alpha^\star$ and θ as indicated above.

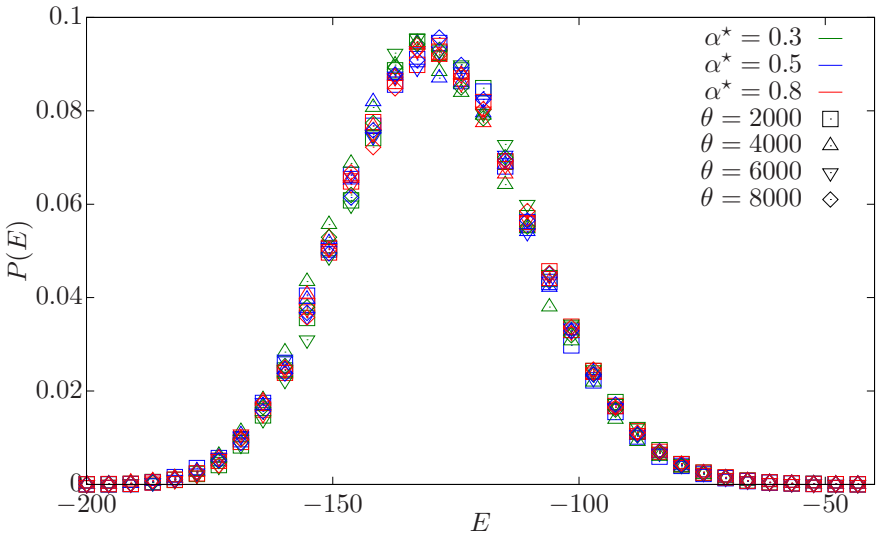

Figure 2. The normalized energy histogram from the simulations with different energy histogram overlaps $\alpha^\star$ and MD simulation steps θ, reweighted to $T = 200$ K using WHAM from the PAMD simulations with adaptive temperature steps. The symbols and colors have the same meaning as in Figs. 1(a) and (b).

As is evident from Fig. 1(a), $\alpha^\star = 0.3$ leads to a temperature set which is very close to the one used in Refs. [2, 10]. Thus the simulation with $\alpha^\star = 0.3$ and $\theta = 4000$ is close to equivalent to our previous study in [10], and it is hence used as a benchmark. In the regime of sufficiently large θ to ensure equilibration in between temperature steps, the adaptive stepping algorithm is insensitive against the chosen θ and mainly depends on $\alpha^\star$. This is also reflected in Fig. 1(b), where we plot the measured temperature T_m against the externally imposed (thermostat) temperature T. Also plotted is a solid line, corresponding to the expected behavior in properly equilibrated simulations, namely $T_m = T$. Note that this type of plot is mostly used in MD simulations, where the intrinsic temperature is readily defined in terms of the kinetic energy, while in MC simulations proper equilibration is mostly ascertained by checking the stationarity of time series of suitably chosen observables.

In Fig. 2 the histogram of potential energy at $T = 200$ K is presented for the different values of $\alpha^\star$ and θ used in Figs. 1(a) and (b). The histograms were obtained by means of the weighted histogram analysis method (WHAM) [17]. One advantage of PAMD simulations is, that one can easily seed the WHAM with good estimates of the partition functions $\mathcal{Z}$, because [5, 7]

$$Q(T, T') = \frac{\mathcal{Z}(T)}{\mathcal{Z}(T')}. \tag{2}$$

This allows one to only perform very few iteration steps in the WHAM while still getting very good estimates of the partition functions. From Fig. 2 it is clear that all considered simulation protocols produce an energy histogram that is well equilibrated.

The quality of sampling of configuration space can be more readily judged from Ramachandran plots, showing the dihedral angles ϕ and ψ along the backbone of the peptide. Such plots are shown in Fig. 3 for the center amino acid GLY3. Panel (a) shows the result of a PT simulation with the same parameters as used in Ref. [10] for reference, panels (b) and (c) correspond to PAMD simulations with adaptive temperature steps using $\alpha^\star = 0.5$, $\theta = 4000$

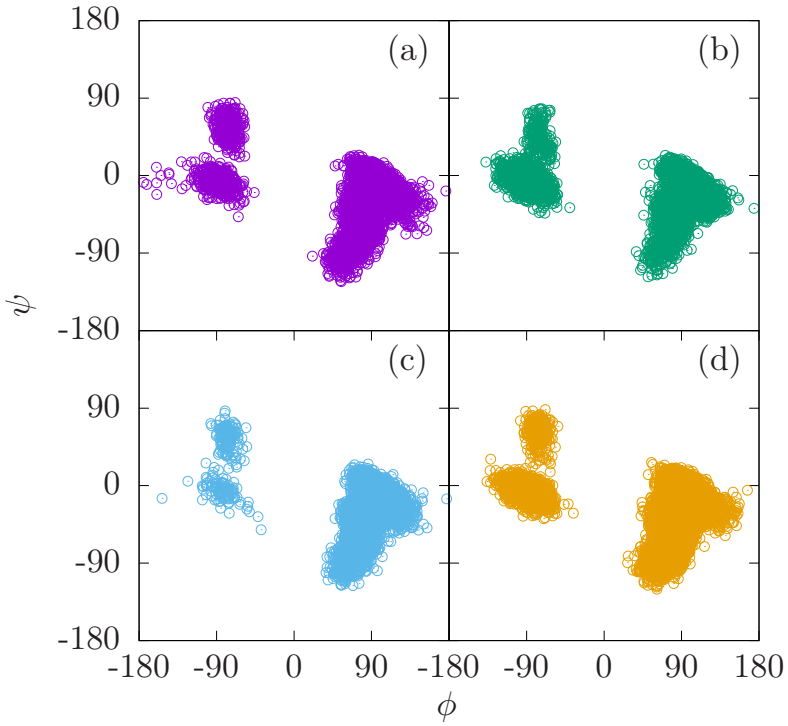

Figure 3. Ramachandran plots showing the realized values of backbone dihedral angles ϕ and ψ for GLY3. (a) The PT simulation with a total runtime of 200 ns already discussed in Ref. [10]. (b) PAMD simulation with adaptive temperature stepping, $\alpha^\star = 0.5$, $\theta = 4000$ and $R = 10\,000$. (c) Same as for (b) but for $\alpha^\star = 0.8$ and $\theta = 2000$. (d) PAMD simulation using the given temperature set from Ref. [2] and $\theta = 4375$, however, with $R = 20\,000$ replicas.

(b) and $\alpha^\star = 0.8$, $\theta = 2000$ (c), respectively. All three simulations show broadly comparable sampling of the configuration space. It appears from a comparison of (b) and (c), that it is favorable to rather choose a larger number of steps θ than increasing the overlap $\alpha^\star$. A similar improvement as in (b), however, is achieved by increasing the population size R, see Fig. 3(d). However, this comparison does not take into account the computational resources used. Relative to the PT simulation shown in (a), the adaptive PAMD simulations in (b) and (c) used 1.525 and 1.875 times more computing cycles, respectively, whereas the non-adaptive but larger PAMD simulation shown in (d) also used 1.875 times the cycles of (a). The sampling of (b) and (d) is in both cases slightly superior to (c). One should keep in mind that the wall-clock time in PAMD with increasing number of replicas R can be kept very low when sufficient parallel resources are available. The minimum wallclock time in PA is predominantly limited only by the short simulation time t_s needed for a single replica

$$t_s = \frac{t}{R}, \tag{3}$$

where t is the total simulation time.

4. Conclusion and Outlook

We have performed population annealing molecular dynamics (PAMD) simulations with adaptive temperature steps for the folding of the penta-peptide met-enkephalin. It was shown that using the criterion of a constant energy overlap significantly simplifies the task of choosing a proper temperature set, a notorious problem also encountered in parallel tempering simulations. We find reliable results in a wide range of possible overlap values $\alpha^\star$ with similar computational efforts if the number θ of MD steps is correspondingly adapted. In the regime where θ is large enough, the precise value of θ has little influence on the actual schedule proposed. If additional parallel computational resources are available, it seems to be preferable to use those to simulate more replicas in the population rather than increasing θ or $\alpha^\star$, but the precise optimum in this multi-dimensional parameter space will depend on the system under consideration.

Acknowledgments

This project was funded by the Deutsche Forschungsgemeinschaft (DFG) under Grant No. SFB/TRR 102 (project B04), and further supported by the Deutsch-Französische Hochschule (DFH-UFA) through the Doctoral College "$\mathbb{L}^4$" under Grant No. CDFA-02-07, the EU Marie Curie IRSES network DIONICOS under Grant No. PIRSES-GA-2013-612707, and the Leipzig Graduate School of Natural Sciences "BuildMoNa".

References

[1] Hansmann U H E 1997 *Chem. Phys. Lett.* **281** 140
[2] Sugita Y and Okamoto Y 1999 *Chem. Phys. Lett.* **314** 141
[3] Iba Y 2001 *Trans. Jpn. Soc. Artif. Intell.* **16** 279
[4] Hukushima K and Iba Y 2003 *AIP Conf. Proc.* **690** 200
[5] Machta J 2010 *Phys. Rev. E* **82** 026704
[6] Wang W, Machta J and Katzgraber H G 2015 *Phys. Rev. E* **92** 063307
[7] Barash L Y, Weigel M, Borovský M, Janke W and Shchur L N 2017 *Comput. Phys. Commun.* **220** 341
[8] Kirkpatrick S, Gelatt C D and Vecchi M P 1983 *Science* **220** 671
[9] Grassberger P 2002 *Comput. Phys. Commun.* **147** 64
[10] Christiansen H, Weigel M and Janke W 2018 *preprint arXiv:1806.06016*
[11] Katzgraber H G, Trebst S, Huse D A and Troyer M 2006 *J. Stat. Mech.: Theory and Exp.* **2006** P03018
[12] Bittner E, Nußbaumer A and Janke W 2008 *Phys. Rev. Lett.* **101** 130603
[13] Vogel T and Perez D 2015 *Phys. Rev. Lett.* **115** 190602
[14] Ferrenberg A M and Swendsen R H 1988 *Phys. Rev. Lett.* **61** 2635
[15] Eastman P, Swails J, Chodera J D, McGibbon R T, Zhao Y, Beauchamp K A, Wang L P, Simmonett A C, Harrigan M P, Stern C D, Wiewiora R P, Brooks B R and Pande V S 2017 *PLoS Comput. Biol.* **13**
[16] Cornell W D, Cieplak P, Bayly C I, Gould I R, Merz K M, Ferguson D M, Spellmeyer D C, Fox T, Caldwell J W and Kollman P A 1995 *J. Am. Chem. Soc.* **117** 5179
[17] Kumar S, Rosenberg J M, Bouzida D, Swendsen R H and Kollman P A 1992 *J. Comput. Chem.* **13** 1011