Resampling improves the efficiency of a “fast-switch” equilibrium sampling protocol

Edward Lyman^{1,a)} and Daniel M. Zuckerman^{2,b)}

¹*Center for Biophysical Modeling and Simulation, University of Utah, 315 S 1400 E, Rm. 2020, Salt Lake City, Utah 84112-0850, USA*

²*Department of Computational Biology, School of Medicine, 3079 BST3, 3501 Fifth Ave., University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA*

(Received 5 September 2008; accepted 9 December 2008; published online 24 February 2009)

We recently applied a multistage reweighting scheme to demonstrate the sampling of equilibrium configurational distributions of peptides from nonequilibrium, simulated annealing trajectories [E. Lyman and D. M. Zuckerman, *J. Chem. Phys.* **127**, 065101 (2007)]. Here we demonstrate that a statistical variance reduction technique, resampling, improves the efficiency of the protocol by about a factor of 3 in a penta-alanine system. While we are not optimistic for the ultimate efficiency of purely temperature-based sampling methods, resampling-type improvements ought to find application in other (formally equivalent) nonequilibrium sampling protocols, such as Jarzynski-relation calculations and annealing-based NMR structure calculations. © 2009 American Institute of Physics. [DOI: [10.1063/1.3081626](https://doi.org/10.1063/1.3081626)]

Reweighting approaches to the problem of configurational sampling—where one distribution is used to sample another—date to the beginning of computer simulation.¹ Recently, we investigated the efficiency of an algorithm originally introduced in the molecular simulation literature² and later rediscovered in the computer science community³ for the sampling of configurations of biomolecules.⁴ The method, “annealed importance sampling” (AIS) as recently implemented for biomolecules relied on high temperature sampling for improved efficiency, and as such achieved only a modest gain in efficiency over conventional, constant temperature simulation.⁵ Here, we show that a commonly used statistical technique for variance reduction—resampling⁶—improves considerably the efficiency of the algorithm. While resampling does not, in our view, make *temperature-based* AIS a promising candidate for sampling configurations of biomolecules we think the idea merits consideration in other contexts, which we discuss at the end.

In the original implementation of AIS, a high temperature sample of configurations is annealed to low temperature. The annealing process consists of instantaneous jumps to a lower temperature, alternating with short constant temperature simulations, just as in ordinary simulated annealing. At each temperature jump in AIS, however, an exponential weight $w_j^i = e^{-\Delta\beta_i E}$ is assigned, where j labels the annealing trajectory, E is the energy of the configuration at the jump, and $\Delta\beta_i = 1/RT_{i-1} - 1/RT_i$, with i indexing the upper temperature. At each stage in a particular annealing trajectory, a weight is assigned to the trajectory by accumulating the product of weights for all preceding temperature jumps. We will refer to these cumulative weights with capital letters, W_j : $W_j = \prod_{i=1}^M w_j^i$, where the temperature ladder is indexed from 0 to M . Equilibrium averages may be calculated at a

temperature T_0 by computing an average over all the configurations annealed to T_0 weighted by the W_j :

$$\langle A \rangle_0 = \frac{\sum_{j=1}^N W_j A_j}{\sum_{j=1}^N W_j},$$

where there are a total of N annealed structures in the sample.

The statistical precision in an AIS estimate of $\langle A \rangle$ now depends on the variance of the weights.^{3,6,7} To see this, consider K independent simulations which generate weights W_{kj} and observable values A_{kj} , where k runs from 1 to K . If the average of all estimates is $\bar{A}$, then the variance of the independent AIS estimates is

$$\langle (\langle A \rangle - \bar{A})^2 \rangle = \frac{1}{K} \sum_{k=1}^K \left(\sum_{j=1}^N \hat{W}_{kj} A_{kj} - \bar{A} \right)^2,$$

where $\hat{W}_{kj} = W_{kj} / \sum_{j=1}^N W_{kj}$ is the normalized weight. This variance depends on the distribution of weights, which can be seen by contrasting two extreme cases: (i) where all weights are equal, as in very slow annealing, and (ii) where one weight is much larger than all others in a given simulation, as with overly rapid annealing. Note further that the above variance includes the effects of correlations which typically are introduced by resampling, as described below, as well as including potential correlations among the weight and the quantity being estimated.

The preceding discussion suggests that AIS can be viewed as a variance reduction scheme, which seeks to alleviate poor weight distributions by introducing intermediate stages of relaxation, during which configurations may *locally* relax. That the constant temperature relaxation steps should be short is implicit in the top-down architecture of AIS: in adopting such a scheme, we are assuming that the effort to generate a good sample of decorrelated structures has been

^{a)}Electronic mail: elyman@hec.utah.edu.

^{b)}Electronic mail: ddmmzz@pitt.edu.

invested in the high T simulation, and now the goal is to propagate that information into a low temperature sample. This was demonstrated for implicitly solvated peptides in our previous paper.⁴ Nevertheless, the problem of significant weight fluctuations also occurs in the AIS scheme and can be expected to become worse for larger systems.

In some situations, poor precision arising from wildly varying weights may be alleviated by “resampling.”⁶ In such a scheme, one uses the weights at intermediate stages of annealing to draw a new, unweighted sample, as suggested in Ref. 4. For example, if we have a sample of N configurations, we choose from this sample N configurations (allowing duplications) with the likelihood of choosing a particular configuration given by its weight. In the following, this is accomplished by simply picking N configurations from the cumulative distribution function (CDF) of the weights, though more sophisticated schemes are possible.⁶ Note that in the new sample some frames will be duplicated—variance reduction comes at the expense of introducing correlations among the annealing trajectories. Also, it is clear that if in the original, weighted sample there are only a handful of configurations with appreciable weight, then only these few configurations will repeatedly be duplicated by the resampling process. Apparently, resampling cannot improve a terrible sample. After resampling, the sample has been “unweighted,” and if we want to compute an average, we do so by calculating a simple arithmetic mean.

In this paper, we have investigated whether resampling can improve the AIS algorithm by improving precision for a fixed annealing schedule, and therefore effectively reducing the expense of the calculation. The original algorithm was modified by introducing a resampling step immediately following each temperature jump. To be specific, consider a set of N configurations which have just been jumped to lower temperature. Before continuing the constant temperature phase, we resample N configurations, duplicating some, discarding others. This new, unweighted sample is then used to spawn the constant temperature leg of the annealing—one constant temperature simulation is started from each configuration. Resampling is performed following each temperature jump.

Figure 1 compares the results of AIS calculations with and without resampling to constant temperature simulation of implicitly solvated dileucine peptide. (The constant temperature simulation protocol is identical to the one used in previous work.⁴) It is clear that, *for this schedule*, resampling dramatically improves the quality of the result. It is very important to note that the improvement should depend on the schedule. For example, if we had used a very slow schedule with long constant temperature stages, the precision would already approach the limit imposed by the initial, high temperature distribution, and there would be no improvement to be gained. There would appear to be more room for improvement in faster, further-from-equilibrium protocols. It ought to also be mentioned that Fig. 1 is strictly speaking not a fair comparison, as resampling introduces some additional overhead. However, this overhead is negligible compared to the cost of the constant temperature simulation.

FIG. 1. (Color online) A comparison of the histogram of potential energies of dileucine peptide at $T=298$ K, as computed by 2000 AIS trajectories with and without resampling. The temperature schedules are identical: five geometrically spaced temperatures, beginning with 500 K down to 298 K, with 0.5 ps of constant temperature relaxation at each level. The distribution of potential energies at the top level (500 K) is shown, along with the distribution at 298 K obtained by constant temperature Langevin dynamics. Note that AIS attains the equilibrium distribution by reweighting a nonequilibrium sample, albeit with much better precision when resampled.

We also investigated the performance of the resampled AIS protocol by “folding” an implicitly solvated alanine pentamer. This system populates an ensemble of diverse unfolded configurations at 400 K, but upon cooling to 250 K the ensemble contains many much more structured configurations, including a significant population of alpha helical conformers. Figure 2 shows that, for equal amounts of simulation time (neglecting again the overhead of building and sampling from the CDF), the resampled AIS protocol yields a significantly more precise estimate of the conformational distribution, here projected along the root mean square deviation (RMSD) from a canonical helix.

We quantified the performance of AIS with resampling in two ways. First, we checked “operationally” for correlations by examining the “ancestry” of surviving configurations. Resampling introduces correlations by duplicating

FIG. 2. (Color online) A comparison of the (un-normalized) histogram of RMSD from a perfect helix for an alanine pentamer calculated by AIS, with and without resampling. The temperature schedules are identical: five geometrically spaced temperatures, beginning with 400 K down to 250 K, with 0.5 ps of constant temperature relaxation at each level. The error bars represent the standard deviation of eight independent runs of 2000 annealing trajectories each. For equal central processing unit time, resampling considerably improves the precision of the calculation.

some annealing trajectories and pruning others, with the potential pitfall that, after many annealing stages, most of the sample could have been spawned by a small number of parent, high temperature structures. This issue was discussed previously by ourselves⁴ and by Neal.³ In our earlier AIS study, we introduced the f statistic, or fraction of structures which contribute to annealed averages, to address the sample size issue. Here, we consider the distribution of trajectories that survive the transit from high to low temperature, since the f value does not simply apply in the resampling case. Among 2000 initial high T configurations, 348 survived to the lowest temperature, which should provide a lower bound to the “effective sample size.” The most frequently resampled initial high T structure spawned, ultimately, 54 low T structures via the branching process of resampling. Twenty structures had at least 20 “children” at the low T , and 57 had at least 10. Note that all of the 54 children of the most prolific structure can only differ in their local features, since they have had at most 1.5 ns of constant temperature simulation since being spawned at high T . This is by design, since the algorithm is structured to invest simulation time in sampling at high temperature and then automatically favor important parts of configuration space through the process of pruning and resampling.

The second quantitative assessment of resampling directly probed the effective sample size—the number of statistically independent configurations—in each sample. Sample sizes were estimated by analyzing the variances of state populations as described elsewhere.⁸ Importantly, the sample size analysis is completely blind to the method by which the samples are generated, and therefore accounts implicitly for correlations introduced by the resampling step. Without resampling, the alanine AIS protocol yielded ~ 100 independent configurations per 2000 annealing trajectories. With resampling, the same protocol yielded ~ 300 independent configurations per 2000 trajectories, in reasonable agreement with the preceding survivorship analysis. Thus, the addition of the resampling step therefore improves the efficiency of AIS by nearly a factor of 3.

We also compared the *sampling rate* of constant temperature Langevin dynamics for the alanine pentamer at 400 and 250 K. Interestingly, the 400 K simulation produced on average ten independent configurations per 1 ns of simulation, while the 250 K simulation produced on average 12 independent configurations per 1 ns. In other words, the high temperature simulation is no more efficient than the low temperature one, within the error of the sample size estimate (approximately 25%). Therefore, for this system, no sampling protocol based on 400 K simulation will be more efficient than standard, constant temperature simulation at $T = 250$ K. However, the conclusion that resampling improves the efficiency of AIS by a factor of 3 is independent of this fact. It may therefore be of use in other reweighting-based protocols, such as Jarzynski-type free energy calculations, which are formally equivalent to AIS, as noted by Neal.³

Our results demonstrate that resampling significantly improves the performance of AIS as applied to model biomolecular systems. Further, the architecture of AIS allows a straightforward estimate of the number of effectively independent structures that are present in the low temperature sample, though some care must be taken to correctly account for correlations introduced by resampling. Another potential advantage of the method is its staged nature, which permits the complete sampling of each temperature before proceeding to the next lower temperature. This opens the possibility of determining the temperature ladder on the fly based on a desired amount of overlap, since the overlap (the number of configurations at the current temperature which contribute to averages at the next temperature down) depends on the known potential energies of the configurations and the temperature jump. Preliminary studies along these lines did not show any improvement over the exponentially distributed temperature ladder used here, perhaps because this model (implicitly solvated penta-alanine) does not have a first-order-like transition as a function of temperature (data not shown).

As we have noted in previous work, there appears to be a modest intrinsic limit to the efficiency obtainable for biomolecular systems using strictly temperature-based methods.^{9,10} This conclusion has been strengthened by the work of others.^{11,12} Therefore, our principal aim in the future is instead to implement an AIS-like approach for the computation of solvated structures from NMR data. NMR structure determination methods already implement very sophisticated and successful annealing protocols. We hope that incorporating an AIS calculation in this context would yield a Boltzmann weighted, native ensemble of solvated structures. Note that this involves annealing in parameters of the target function in addition to temperature.

We thank Radford Neal for comments on our previous work and Andrew Petersen for reading the present report. Support for this work was provided by the NSF (Grant No. MCB-0643456) and the NIH (GM076569).

¹H. Kahn, in *Proceedings of the Seminar on Scientific Computation*, 1949 (IBM, New York, 1950), pp. 20–27.

²G. A. Huber and J. A. McCammon, *Phys. Rev. E* **55**, 4822 (1997).

³R. M. Neal, *Stat. Comput.* **11**, 125 (2001).

⁴E. Lyman and D. M. Zuckerman, *J. Chem. Phys.* **127**, 065101 (2007).

⁵D. M. Zuckerman and E. Lyman, *J. Chem. Theory Comput.* **2**, 1200 (2006).

⁶J. S. Liu, *Monte Carlo Strategies in Scientific Computing* (Springer, New York, 2001).

⁷T. Shen and D. Hamelberg, *J. Chem. Phys.* **129**, 034103 (2008).

⁸E. Lyman and D. M. Zuckerman, *J. Phys. Chem. B* **111**, 12876 (2007).

⁹D. M. Zuckerman and E. Lyman, *J. Chem. Theory Comput.* **2**, 1200 (2006).

¹⁰D. M. Zuckerman and E. Lyman, *J. Chem. Theory Comput.* **2**, 1693 (2006).

¹¹H. Nymeyer, *J. Chem. Theory Comput.* **4**, 626 (2008).

¹²R. Denschlag, M. Lingenheil, and P. Tavan, *Chem. Phys. Lett.* **458**, 244 (2008).