

Error and efficiency of simulated tempering simulations

Edina Rosta and Gerhard Hummer^{a)}*Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Building 5, Bethesda, Maryland 20892-0520, USA*

(Received 20 November 2009; accepted 16 December 2009; published online 15 January 2010)

We derive simple analytical expressions for the error and computational efficiency of simulated tempering (ST) simulations. The theory applies to the important case of systems whose dynamics at long times is dominated by the slow interconversion between two metastable states. An extension to the multistate case is described. We show that the relative gain in efficiency of ST simulations over regular molecular dynamics (MD) or Monte Carlo (MC) simulations is given by the ratio of their reactive fluxes, i.e., the number of transitions between the two states summed over all ST temperatures divided by the number of transitions at the single temperature of the MD or MC simulation. This relation for the efficiency is derived for the limit in which changes in the ST temperature are fast compared to the two-state transitions. In this limit, ST is most efficient. Our expression for the maximum efficiency gain of ST simulations is essentially identical to the corresponding expression derived by us for replica exchange MD and MC simulations [E. Rosta and G. Hummer, *J. Chem. Phys.* **131**, 165102 (2009)] on a different route. We find quantitative agreement between predicted and observed efficiency gains in a test against ST and replica exchange MC simulations of a two-dimensional Ising model. Based on the efficiency formula, we provide recommendations for the optimal choice of ST simulation parameters, in particular, the range and number of temperatures, and the frequency of attempted temperature changes. [doi:10.1063/1.3290767]

I. INTRODUCTION

Molecular dynamics (MD) and Monte Carlo (MC) simulations are powerful tools to explore the configuration space of molecular systems. However, for many systems the sampling efficiency is limited by the slow rate of interconversion between different conformational basins on the energy surface. To accelerate the conformational sampling, a variety of extended ensemble schemes have been developed. Among the most widely used ones are simulated tempering (ST),^{1,2} replica exchange Monte Carlo (REMC),^{3,4} and replica exchange molecular dynamics (REMD).^{5,6} Replica exchange is also known as parallel tempering. These extended ensemble methods aim to establish an equilibrium between canonical systems at different temperatures (or, more generally, systems with different Hamiltonians⁷), an idea that can be traced back to the seminal papers by Bennett,⁸ and Swendsen and Wang.⁹ By coupling simulations at a low temperature of interest with simulations at high temperatures where the relaxation is fast, one hopes to transfer the improved sampling efficiency at the higher temperature down to the lower temperature.

In ST simulations,^{1,2} the reference thermostat temperature is itself a dynamic variable. At fixed intervals along an otherwise regular MD or MC simulation, attempts are made to change the current reference temperature T_i to a trial temperature T_j picked from a discrete set (e.g., $T_{i\pm 1}$ in a series of temperatures T_1, T_2, \dots, T_N). These attempted temperature changes are accepted with a criterion designed to maintain

detailed balance for canonical distributions. In a long run (assuming mixing in the resulting Markov chain), canonical distributions will be obtained at each of the temperatures T_i . In REMD (Refs. 5 and 6) and REMC,^{3,4} MD and MC simulations of N identical molecular systems are performed in parallel. Each replica is thermostatted at a different temperature T_i to establish corresponding canonical distributions.¹⁰ With a given frequency, one attempts to exchange the configurations of replica pairs, $i \leftrightarrow j$ (typically $i \leftrightarrow i \pm 1$). These replica exchanges are accepted with a probability that conserves the respective canonical distributions. ST and REMD have been proven particularly useful to enhance the sampling of biomolecular systems.^{5,6,11–15}

Key questions in ST, REMD, REMC, or other extended ensemble simulations are as follows: (1) How large is the possible gain in computational efficiency, (2) for which simulation parameters can this maximum gain be realized, and (3) how do the different methods compare with respect to their efficiency gains? We have previously developed a simple analytical formula for the efficiency gain in REMD and REMC simulations¹⁶ for the important case that the slow dynamics of the system of interest can be described by two-state transitions. Here, we derive a formula for the error and computational efficiency of ST simulations. The formula applies to systems with two-state dynamics at long-time scales (with an extension to the multistate case being discussed). For the sake of concreteness, we will frequently refer to two-state protein folding, but the theory is general. We measure the error in the estimator of an equilibrium property by its variance over repeated simulations. The computational effi-

^{a)}Electronic mail gerhard.hummer@nih.gov.

ciency is then defined as the rate with which the variance in the estimator decreases with length of the simulations.

The main result of this paper is that the relative efficiency of a ST simulation compared to a regular MD (or MC) simulation of the same duration t_{sim} , run at the temperature T_m of interest alone, is

$$\eta_m = \frac{\sigma_{\text{MD}}^2}{\sigma_{\text{ST}}^2} = \sum_{i=1}^N Q_i \frac{\tau_m^+ + \tau_m^-}{\tau_i^+ + \tau_i^-} \quad (1)$$

independent of the property of interest. In the ST simulation, the system spends a fraction Q_i at each of the N temperatures T_1, T_2, \dots, T_N , with $\sum_{i=1}^N Q_i = 1$. Equation (1) is valid for systems whose slow dynamics can be described by two-state transitions, and in the asymptotic limit of long simulation times t_{sim} . The lifetimes of the system in its two long-lived (unfolded and folded) states at temperature T_i are $\tau_i^+ \equiv 1/k_i^+$ and $\tau_i^- \equiv 1/k_i^-$, where k_i^+ and k_i^- are the corresponding (folding and unfolding) rates. A ST simulation is more efficient than a MD simulation of the same duration at temperature T_m if $\eta_m > 1$.

We can interpret Eq. (1) for the relative efficiency of ST simulations as the ratio of the net reactive fluxes given by the number of folding and unfolding transitions per unit time in the ST simulation, and in the MD simulation: $2t_{\text{sim}}/(\tau_m^+ + \tau_m^-)$ is the expected number of two-state transitions in MD and $\sum_i 2Q_i t_{\text{sim}}/(\tau_i^+ + \tau_i^-)$ in the ST simulation. Remarkably, Eq. (1) is essentially identical to the efficiency formula obtained by us for REMD (Ref. 16) on a different, more involved route. We will show that this is no coincidence: The derivation followed here for ST can be adapted for replica exchange. Practically, the equivalence of the efficiency formulas means that at least asymptotically for two-state systems and ignoring any issues of computational overhead, ST and REMD have the same efficiency.

To illustrate and test the efficiency formula, we perform MC simulations of a two-dimensional (2D) Ising system at zero magnetic field. Below the critical temperature of the infinite system, a finite spin system will only slowly interconvert between the ‘‘up’’ and ‘‘down’’ states of the net magnetization, if only single spin flips are attempted in the MC sampling. We show that the slow sampling of the net magnetization of a finite system is accurately captured by our theory. We also show that Eq. (1) quantitatively reproduces the observed gains in efficiency from using ST, and that the large predicted gains in efficiency can indeed be realized for the 2D Ising system with its thermally activated transition in magnetization. In conclusion, we discuss procedures to accelerate the sampling of temperature space and strategies to optimize the choice of parameters.

II. THEORY

A. Rate model of simulated tempering

Building on our recent work on REMD,¹⁰ we will analyze the statistical error and efficiency of ST simulations for the important case of systems whose dynamics at long times is governed by a single slow exponential relaxation process. Even though complex molecular systems normally have

multiple states, and thus a broad spectrum of relaxation processes, at sufficiently long times the relaxation to equilibrium is often dominated by a single slow exponential process. The folding of small ‘‘two-state’’ proteins is a specific example, with measured relaxation times that range from microseconds to seconds.

For such systems, a reduced two-state description captures the long-time dynamics. For the sake of concreteness, here we will refer to protein folding and unfolding as the slow processes, but the results are general. In the following, we assume that the interconversion between the same two states U and F dominate the relaxation at low and high temperatures (but we note that this may not always be the case). In the absence of ST temperature changes, we assume that the interconversion between the folded and unfolded states F_i and U_i , respectively, at each of the temperatures T_i follows first-order kinetics,



At temperature T_i , the equilibrium folded and unfolded populations are $p_i = k_i^+ / (k_i^+ + k_i^-)$ and $q_i = 1 - p_i$, respectively.

In ST simulations,^{1,2} one periodically attempts to change the simulation temperature from T_i to T_j . These changes are accepted with a probability that maintains detailed balance with respect to a canonical ensemble at each temperature. If $E = E(\mathbf{x})$ is the energy of a configuration \mathbf{x} , a change from temperature T_i to T_j is accepted with probability

$$P_{\text{acc}}(T_i \rightarrow T_j) = \min(1, e^{E/k_B T_i - E/k_B T_j + f_j - f_i}), \quad (3)$$

where k_B is Boltzmann’s constant. Here we assume that the generation probabilities for changes $T_i \rightarrow T_j$ and $T_j \rightarrow T_i$ are identical. If \mathbf{x} is a purely configurational coordinate of the system, as in typical MC sampling, E is the potential energy; in contrast, if \mathbf{x} is a position in phase space, as in MD simulations, E is the total energy. Alternatively, in ST combined with MD simulations, one can also use the potential energy E in the acceptance criterion, combined with either rescaling momenta or redrawing them from a Maxwell–Boltzmann distribution at the new temperature T_j upon acceptance. The f_i are constants that determine the fraction Q_i of time spent at temperature T_i in long runs,

$$Q_i = \frac{e^{f_i} Z_i}{\sum_{j=1}^N e^{f_j} Z_j}, \quad (4)$$

where $Z_i = \int d\mathbf{x} e^{-E(\mathbf{x})/k_B T_i}$ is the partition function at T_i and $0 < Q_i \leq 1$ with $\sum_{i=1}^N Q_i = 1$. Note that otherwise the Q_i are arbitrary, determined by the specific choice of ST simulation parameters. In practice, the coefficients f_i can be gradually optimized to achieve a targeted distribution, e.g., $Q_i = 1/N$. These optimization strategies are typically based on the average energies at the different temperatures.^{2,17} Here, we will instead solve the $N-1$ coupled linear equations

$$Q_i \sum_{j=1}^N Q_j^{\text{obs}} g_j = Q_i^{\text{obs}} g_i \quad (5)$$

for $g_i = \exp(f_i^{\text{new}} - f_i^{\text{old}})$ with $g_1 = 1$, Q_i^{obs} is the observed fraction at T_i in a preliminary run with weight factors f_i^{old} , and f_i^{new} are the new weight factors for subsequent runs.

To model ST with a kinetic model with $2N$ states, we allow changes in the temperature along the simulation trajectory. Here, we focus on the case of very frequent temperature change attempts (which will result in the highest possible efficiency gains). In this limit of fast ST, we can use a rate model also for changes in the temperature in ST, not just for the two-state transitions of the system,

$$F_i \xrightleftharpoons[r_{i+1}^-]{r_i^+} F_{i+1} \quad \text{and} \quad U_i \xrightleftharpoons[s_{i+1}^-]{s_i^+} U_{i+1} \quad (i = 1, 2, \dots, N-1). \quad (6)$$

In the kinetic model, the rates r_i^+ and r_{i+1}^- (s_i^+ and s_{i+1}^-) for temperature changes in a folded (unfolded) structure should satisfy detailed balance to maintain the proper equilibria,

$$\begin{aligned} r_i^+ &= r(Q_{i+1}p_{i+1}/Q_i p_i)^{1/2}, \\ r_{i+1}^- &= r(Q_i p_i/Q_{i+1} p_{i+1})^{1/2}, \\ s_i^+ &= r(Q_{i+1}q_{i+1}/Q_i q_i)^{1/2}, \\ s_{i+1}^- &= r(Q_i q_i/Q_{i+1} q_{i+1})^{1/2}, \end{aligned} \quad (7)$$

where r is a constant rate that will be chosen large enough to ensure frequent temperature changes. To model an actual simulation, the rate coefficients can be estimated from the acceptance probabilities p_{acc} , given the folding state of the system, and the time intervals Δt between attempted changes of a certain type, e.g., $s_i^+ = p_{\text{acc}}(i \rightarrow i+1|U)/\Delta t(i \rightarrow i+1)$. In our ST implementations, we will also consider non-nearest-neighbor transitions, for which temperature change rates can be estimated by analogy. Similar master-equation descriptions resulting in simple first-order kinetic models have previously been used to model replica exchange simulations.^{16,18-21}

B. Error and efficiency of MD and ST simulations

Calculating equilibrium properties is the main goal in both MD, MC, and ST simulations. We accordingly define the computational efficiency of the simulations as the rate with which the statistical error in the estimate of the property of interest decreases with the simulation time. Following our earlier work on the efficiency of REMD simulations,¹⁶ we note that for simulation times t_{sim} long compared to the overall longest relaxation time of the system, the central limit theorem implies that the error in the estimate \bar{A} of the exact mean $\langle A \rangle$ of any property A decreases as $\text{var}(\bar{A}) = c/t_{\text{sim}}$, where $\text{var}(\bar{A})$ indicates the variance about the true mean in multiple simulations of the same duration t_{sim} . The constant c depends on the simulation method (MD versus ST). Note that we do not consider the approach of \bar{A} to equilibrium for a given (“nonequilibrium”) initial condition. This approach

to the true mean $\langle A \rangle$ occurs asymptotically as $1/t_{\text{sim}}$, faster than the decrease in the statistical error ($t_{\text{sim}}^{-1/2}$).

The ratio of the variances of the estimators from MD and ST simulation methods allows us to compare their respective efficiencies. Even though we explicitly consider only the case of two-state protein folding, the results apply generally to systems with two metastable states that interconvert slowly compared to the relaxation processes within each state. Importantly,¹⁶ in the two-state case the relative computational efficiencies will not depend on the particular property A . In such systems, A will quickly relax to the average values $\langle A \rangle_F$ and $\langle A \rangle_U$ in folded and unfolded states F and U , respectively. If those averages differ, then $\text{var}(\bar{A}) \approx (\langle A \rangle_F - \langle A \rangle_U)^2 \text{var}(\bar{s})$, where s indicates the folding state with $s=1$ if the system is folded, and $s=0$ if it is unfolded, such that $\langle s \rangle_i = p_i$ at temperature T_i . It thus suffices to compare the convergence of estimates of the fraction folded, $\langle s \rangle_1$, at the temperature T_1 of interest.

For regular MD (and, by analogy, MC) simulations, one finds either directly¹⁶ or from the theory of statistical errors in single-molecule experiments^{22,23} that the estimator \bar{s} , defined as the fraction of time spent in the folded state in a simulation of duration t_{sim} has a variance of

$$\sigma_{\text{MD}}^2(t_{\text{sim}}) \equiv \text{var}(\bar{s}) = \frac{2}{t_{\text{sim}}} \frac{p_1(1-p_1)}{k_1^+ + k_1^-} = \frac{2}{t_{\text{sim}}} \frac{k_1^+ k_1^-}{(k_1^+ + k_1^-)^3} \quad (8)$$

at the temperature T_1 of interest. Equation (8) applies in the asymptotic limit of $t_{\text{sim}} \gg (k_1^+ + k_1^-)^{-1}$.

For ST simulations, we consider the limit of very frequent temperature changes (r large). In this fast ST limit, the different temperatures will be visited frequently before a change in the folding state occurs at any one of the temperatures T_i . We can thus describe the kinetics of the folding state with effective folding and unfolding rates,

$$\begin{aligned} k_{\text{eff}}^+ &= \sum_{i=1}^N k_i^+ p(i|U), \\ k_{\text{eff}}^- &= \sum_{i=1}^N k_i^- p(i|F), \end{aligned} \quad (9)$$

where $p(i|U)$ and $p(i|F)$ are the probabilities of being at temperature T_i , given that the system is unfolded and folded, respectively. From Bayes' theorem, $p(i|U) = p(U|i)p(i)/p(U)$ with $p(U|i) = q_i$, $p(i) = Q_i$, and $p(U) = \sum_{i=1}^N Q_i q_i$ [and analogous relations for $p(i|F)$] it follows that

$$k_{\text{eff}}^+ = \kappa/(1-p), \quad (10)$$

$$k_{\text{eff}}^- = \kappa/p,$$

where

$$\kappa = \sum_{i=1}^N \frac{Q_i}{1/k_i^+ + 1/k_i^-} = \sum_{i=1}^N \frac{Q_i}{\tau_i^+ + \tau_i^-} \quad (11)$$

and $p = p(F) = k_{\text{eff}}^+ / (k_{\text{eff}}^+ + k_{\text{eff}}^-)$ is the net fraction folded irrespective of temperature. In the last expression, we substituted the lifetimes $\tau_i^\pm = 1/k_i^\pm$ for the rates of interconversion.

To estimate the sampling error, we define the estimator x_1 of $p_1 \equiv \langle s \rangle_1$ (i.e., the fraction folded at the temperature T_1 of interest) as $x_1 = t_{F1} / (t_{F1} + t_{U1})$, where t_{F1} and t_{U1} are the times spent folded and unfolded, respectively, at temperature T_1 during a simulation of duration t_{sim} . In the following, we will calculate the variance $\sigma_{\text{ST}}^2 \equiv \text{var}(x_1) \equiv \text{var}(\bar{s})$ of x_1 over repeated ST simulations of duration t_{sim} .

In the limit of fast ST, we have effective two-state behavior with folding and unfolding rates k_{eff}^+ and k_{eff}^- , respectively. We define t_F and $t_U = t_{\text{sim}} - t_F$ as the total time folded and unfolded in the simulation, irrespective of the temperature. Then, in the limit of fast ST, we have $t_{F1} = p(1|F)t_F$ and $t_{U1} = p(1|U)t_U$, i.e., perfect sampling of the fractions of time spent at temperature T_1 for given times t_F and t_U spent folded and unfolded, respectively,

$$x_1 \equiv \frac{t_{F1}}{t_{F1} + t_{U1}} = \frac{t_F p(1|F)}{t_F p(1|F) + (t_{\text{sim}} - t_F) p(1|U)}. \quad (12)$$

We further assume that the simulation time t_{sim} is sufficiently long compared to the characteristic time $1/k_{\text{eff}}$, where $k_{\text{eff}} = k_{\text{eff}}^+ + k_{\text{eff}}^-$ is the effective folding relaxation rate irrespective of temperature. We can then use the asymptotic formula^{22,23} for the variance of t_F and linear error propagation,

$$\text{var}(x_1) = \text{var}(t_F) \left(\frac{dx_1}{dt_F} \Big|_{t_F = \langle t_F \rangle} \right)^2, \quad (13)$$

where^{22,23}

$$\text{var}(t_F) = \frac{2t_{\text{sim}}p(1-p)}{k_{\text{eff}}} = \frac{2t_{\text{sim}}p^2(1-p)^2}{\kappa}. \quad (14)$$

After differentiating Eq. (12) with respect to t_F , and substituting $p(1|F)p + p(1|U)(1-p) = p(1) = Q_1$ and $\langle t_F \rangle = pt_{\text{sim}}$, we find that

$$\frac{dx_1}{dt_F} \Big|_{t_F = \langle t_F \rangle} = \frac{k_1^+ k_1^-}{t_{\text{sim}} p (1-p) (k_1^+ + k_1^-)^2}. \quad (15)$$

Substituting Eqs. (14) and (15) into Eq. (13) results in

$$\text{var}(x_1) = \frac{2}{t_{\text{sim}} \kappa} \frac{(k_1^+ k_1^-)^2}{(k_1^+ + k_1^-)^4}. \quad (16)$$

Comparing this expression to the corresponding variance of the estimator in a regular MD simulation at temperature T_1 , Eq. (8), we find that

$$\text{var}(x_1) \equiv \sigma_{\text{ST}}^2 = \frac{\sigma_{\text{MD}}^2 k_1^+ k_1^-}{(k_1^+ + k_1^-) \kappa} = \frac{\sigma_{\text{MD}}^2}{\sum_{i=1}^N Q_i \frac{\tau_1^+ + \tau_1^-}{\tau_i^+ + \tau_i^-}}. \quad (17)$$

In the last expression, we used Eq. (11) for κ and $k_1^\pm = 1/\tau_1^\pm$ for the rate coefficients. By rewriting Eq. (17) in terms of the ratio of variances $\sigma_{\text{MD}}^2/\sigma_{\text{ST}}^2$ for a more general temperature of interest, T_m instead of T_1 , we obtain our main result, Eq. (1), for the relative efficiency of ST simulations compared to MD simulations using the same computational resources.

We note that one obtains the same result by using error propagation, including covariances. Specifically, if one de-

finies covariances of the times spent folded and unfolded irrespective of temperature as $C_{\alpha\beta} = \langle t_\alpha t_\beta \rangle - \langle t_\alpha \rangle \langle t_\beta \rangle$ with $\alpha, \beta \in \{F, U\}$, then

$$\text{var}(x_1) = \sum_{\alpha, \beta} \frac{\partial x_1}{\partial t_\alpha} C_{\alpha\beta} \frac{\partial x_1}{\partial t_\beta}. \quad (18)$$

With this procedure it is possible to extend the efficiency analysis worked out above for two-state systems to systems whose slow dynamics involves $M > 2$ states. In the limit of fast ST, one can again determine effective transition rates $K_{\alpha\beta}^{\text{eff}} = \sum_i K_{\alpha\beta} p(i|\beta)$ from state β to α , averaged over the different temperatures T_i , analogous to Eq. (9). From the resulting effective rate matrix \mathbf{K}^{eff} , the matrix of correlation coefficients,

$$C_{\alpha\beta} = \langle t_\alpha t_\beta \rangle - \langle t_\alpha \rangle \langle t_\beta \rangle, \quad (19)$$

of the times t_α and t_β spent in states α, β and β irrespective of temperature ($\alpha = 1, 2, \dots, M$) can be determined in the asymptotic limit of large t_{sim} as

$$\mathbf{C} = 2t_{\text{sim}} [a^{-1} \mathbf{P} - (\mathbf{K}^{\text{eff}} + a\mathbf{P})^{-1}] \mathbf{P}', \quad (20)$$

where $\mathbf{P} = (p_1, p_2, \dots, p_N)^T (1, 1, \dots, 1)$ is a projector onto the equilibrium distribution, \mathbf{P}' is a diagonal matrix, $P'_{\alpha\beta} = \delta_{\alpha\beta} p_\beta$, and a is an arbitrary nonzero number with dimensions of reciprocal time. Equation (20) corresponds to Eq. (12) in Ref. 24, noting that $\tau_{\alpha\beta}$ in that paper is related to $C_{\alpha\beta}$ through $C_{\alpha\beta} = -2t_{\text{sim}} p_\alpha \tau_{\alpha\beta} p_\beta$ for $\alpha \neq \beta$ and $C_{\alpha\alpha} = 2t_{\text{sim}} p_\alpha \tau_{\alpha\alpha} (1 - p_\alpha)$. If we define $x_{\alpha 1} = t_{\alpha 1} / \sum_{\beta=1}^M t_{\beta 1}$ as the estimator of the relative population $p(\alpha|T_1)$ in state α at temperature T_1 , then

$$\frac{\partial x_{\alpha 1}}{\partial t_\beta} = \frac{1}{t_{\text{sim}}} \left[\delta_{\alpha\beta} \frac{p(\alpha|T_1)}{p(\alpha)} - \frac{p(\alpha|T_1)p(\beta|T_1)}{p(\beta)} \right], \quad (21)$$

where $p(\alpha)$ is the fraction in state α averaged over all temperatures. Substitution of Eqs. (20) and (21) into Eq. (18) allows us to evaluate the variance of the estimator $x_{\alpha 1}$ for a system with more than two states.

C. Relation to REMD and REMC

Equation (1) is identical to our previous result for the relative efficiency of REMD simulations¹⁶ if we identify $Q_i = 1/N$. In REMD simulations, the different temperatures indeed have equal weights by construction. As a consequence, REMD and ST have the same computational efficiency for the systems studied here in the asymptotic limit of long simulations, and if the sampling of temperature space is uniform in ST. However, the methods are only equivalent in the limits of fast temperature changes in ST, and fast replica exchanges in REMC and REMD simulations (and for a large number of replicas).

The efficiency result for REMD simulations was derived by us on a very different route. Specifically, in Ref. 16 we studied the global relaxation dynamics of the total number $n(t)$ of folded replicas at time t , requiring us to estimate the first nonzero eigenvalue of a large rate matrix. In the following, we sketch a rederivation of the efficiency formula for REMD (and, by analogy, REMC) simulations following the route in the present paper.

In REMD and REMC simulations, every replica is statistically equivalent.²⁵ So if we follow each replica, as it travels through temperature space, we have in effect the same dynamics as we have here for a single ST run (except that in REMC and REMD, $Q_i=1/N$ is fixed). In particular, if temperature changes are fast compared to folding and unfolding, the folding dynamics can be described with the effective rates in Eq. (10). As a consequence, in a REMD simulation of duration t_{sim}/N (which is N times shorter than a MD or ST simulation using the same resources), the variance in the estimator x from an individual replica will be N times larger than that in Eq. (17). However, in the limit of a large number of replicas, the estimators from different replicas are statistically independent. As a consequence of the central limit theorem, the variance in the estimator x obtained by pooling all N replicas is thus reduced by a factor N , and we arrive at Eq. (17) also for REMD, with $Q_i=1/N$. We note, however, that in the case of replica exchange the efficiency formula is only approximate (because of the assumption of statistical independence of the estimators), even in the limit of fast replica exchange, whereas it is exact for ST in the fast limit. We note further that in REMC and REMD, one gets weights Q_i different from $1/N$ simply by running more replicas at a certain temperature.

As an interesting note, we showed in Ref. 16 that the estimators of all typical observables in a REMD simulation (and, in particular, the total number of folded replicas) relax at long times with a characteristic rate $\sum_{i=1}^N (k_i^+ + k_i^-) q_i p_i / \sum_{i=1}^N q_i p_i$. This rate is different from the relaxation rate predicted for ST, $k_{\text{eff}} = k_{\text{eff}}^+ + k_{\text{eff}}^-$. However, k_{eff} should give the relaxation of the estimator evaluated for an individual replica in REMD (or REMC), as it performs its random walk through temperature space.

III. RESULTS

A. Error and efficiency in simulated tempering

To test Eq. (1) for the efficiency, we have performed MC simulations of a 2D Ising model with Hamiltonian,

$$H = \sum_{(i,j)} \sigma_i \sigma_j. \quad (22)$$

The sum extends over all distinct pairs (i,j) of neighboring spins $\sigma_i = \pm 1$ on a periodic square lattice of size $K \times K$. In a canonical ensemble with Boltzmann factor $\exp(-H/T)$ and partition function $\sum_{\{\sigma_i\}} \exp(-H/T)$ (with $k_B=1$ in this section), in the absence of an external magnetic field, and below the critical temperature $T_c = 2/\ln(1+2^{1/2}) \approx 2.3$ for the infinite system, finite 2D Ising models exhibit bistability with respect to their reduced magnetization $m = \sum_i \sigma_i / K^2$. At sufficiently low temperatures, m fluctuates between $m \approx -1$ and 1.

The dynamics of this interconversion between up and down states of m will depend on the chosen MC move set. If in the MC sampling the move set is restricted to flips of single spins, the relaxation time of m will be long, measured in units of MC steps (one attempted move per spin) or MC

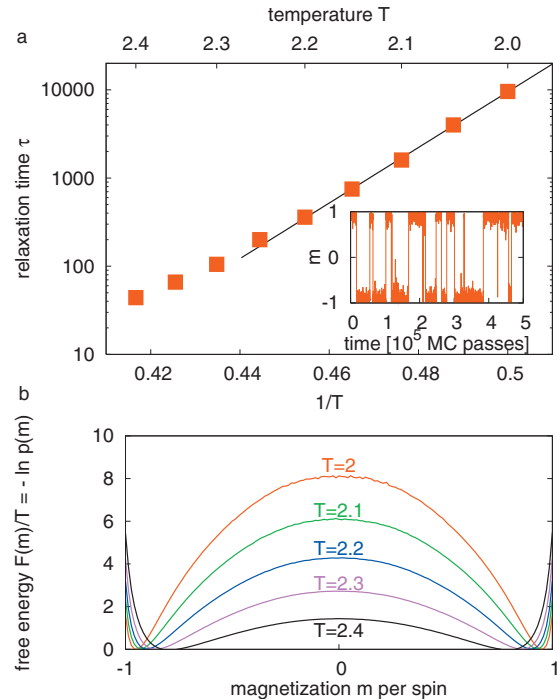


FIG. 1. Magnetization of the 12×12 2D Ising model. (a) Arrhenius plot of relaxation time $\tau(T)$ of magnetization m in units of MC passes vs $1/T$. Symbols: MC data; line: Arrhenius fit. Inset: MC trajectory segment of the reduced magnetization $m(t)$ at $T=2$. (b) Free energy profile $F(m)/T = -\ln p(m)$ of the reduced magnetization m at temperatures $T=2, 2.1, \dots, 2.4$ (top to bottom at $m=0$).

passes (one attempted move for each of the K^2 Ising spins). The reason is that intermediate states with $m \approx 0$ are energetically unfavorable.

The inset in Fig. 1(a) shows the magnetization $m(t)$ for a trajectory segment at $T=2$ of a 2D periodic Ising systems with 12×12 spins obtained from a regular MC simulation with single spin flips. From the decay of the autocorrelation function of $m(t)$ at temperatures between 2 and 2.45 we have determined the relaxation time $\tau(T)$. The Arrhenius plot of $\tau(T)$ in Fig. 1(a) shows that the relaxation time grows exponentially with $1/T$ at low temperatures. As a consequence, sampling of $\langle m \rangle$ with naive MC is challenging at low temperatures. The slow interconversion rate of the up and down states of m at low temperatures is also reflected in a high barrier at $m=0$ in the free energy profile $F(m)/T = -\ln p(m)$, where $p(m)$ is the probability density of m [Fig. 1(b)].

In addition to the regular MC simulations, we have also performed ST simulations at temperatures $T_i = 2 + (i-1)/20$, $i=1, 2, \dots$. We have run simulations for $N=2$ ($T_1=2$, $T_2=2.05$), $N=3$ ($T_1=2$, $T_2=2.05$, $T_3=2.1$), etc., up to $N=10$ (with $T_1=2$ and $T_{10}=2.45$). The weight factors f_i determining the relative populations Q_i of the different temperatures in the ST runs are adjusted by solving Eq. (5) such that $Q_i=1/N$. After every MC pass, we randomly pick a trial temperature T_j ($j=1, 2, \dots, i-1, i+1, \dots, N$) different from the current temperature T_i , and accept the change from T_i to T_j according to Eq. (3).

From each of the simulations, we estimated the average magnetization \bar{m} at the lowest temperature, $T=2$, as the simple mean of the instantaneous magnetization along the

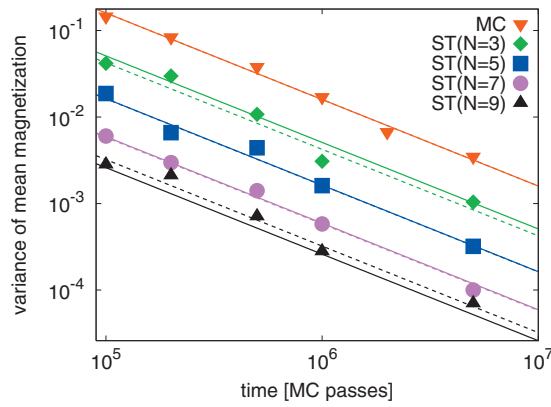


FIG. 2. Variance in the estimator of the mean magnetization as a function of the simulation length t_{sim} measured in units of MC passes. The symbols show the results of a regular MC simulation (top, inverted triangles) and ST simulations with $N=3, 5, 7$, and 9 temperatures (top to bottom) and corresponding upper temperatures $T_N=2.1, 2.2, 2.3$, and 2.4 . The solid lines are the theoretical predictions, c_N/t_{sim} , with $c_N = (\langle m \rangle_+ - \langle m \rangle_-)^2 \text{var}(x_1) = (2|m|)^2 \text{var}(x_1)$ and $\text{var}(x_1) \equiv \text{var}(\bar{s})$ defined in Eq. (8) for the MC data ($N=1$) and in Eq. (17) for the ST data ($N>1$). The dashed lines are fits to c_N/t_{sim} . Note that for $N=9$, the actual error is slightly larger than predicted.

trajectory. In the limit of perfect sampling, \bar{m} converges to the exact limit $\langle m \rangle = 0$. From multiple ST simulations of the same duration (number of MC passes between $t_{\text{sim}}=10^5$ and 5×10^6) we have determined the variance of the estimator \bar{m} and confirmed that it scales as $\sigma^2(\bar{m}) = c_N/t_{\text{sim}}$ for both regular MC ($N=1$) and ST simulations ($N>1$). The coefficients c_N are determined from a fit, with c_1 being the coefficient that determines the rate of convergence for regular MC, c_2 for ST with two temperatures, $T_1=2$ and $T_2=2.05$, etc. As shown in Fig. 2, the fitted coefficients c_N are in excellent agreement with the theoretical predictions from Eq. (8) for regular MC, and Eq. (17) for the ST simulations.

Figure 3 plots the relative computational efficiencies $\eta_1 = c_1/c_N$ of the ST simulations compared to a regular MC simulation at temperature $T=2$, with the c_i obtained from fits. Also plotted is the relative efficiency gain η_1 predicted from Eq. (1) using the life times $\tau_i^+ + \tau_i^- = 4\tau(T_i)$, where $\tau(T_i)$

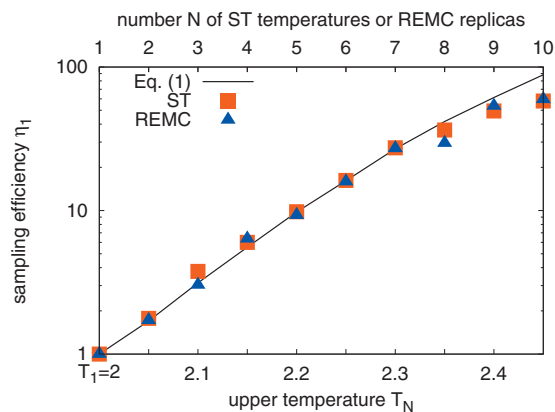


FIG. 3. Efficiency gain η_1 in ST (squares) and REMC simulations (triangles) over regular MC sampling of the magnetization m of the finite 2D Ising model at $T_1=2$. The theoretical prediction, Eq. (1), is shown as a line. The bottom axis indicates the upper temperature T_N in the ST and REMC simulations. The top axis indicates the corresponding number of evenly spaced temperatures in the ST simulations, and replicas in the REMC simulations.

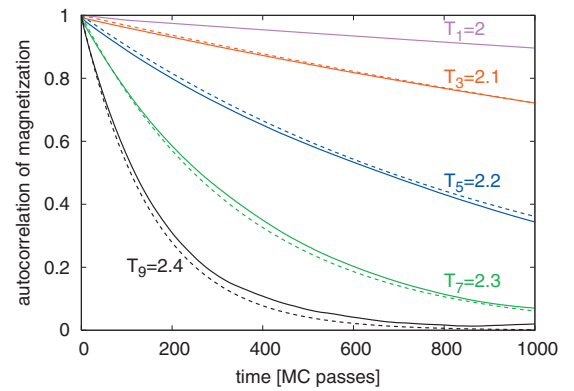


FIG. 4. Normalized autocorrelation function $C(t)$ of the magnetization $m(t)$ irrespective of temperature in ST simulations (solid lines) with $N=3, 5, 7$, and 9 evenly spaced temperatures and upper temperatures of $T_N=2.1, 2.2, 2.3$, and 2.4 . Dashed lines show the theoretical predictions $C(t) = \exp(-k_{\text{eff}}t)$. For reference, the top curve shows $C(t)$ for a regular MC run at temperature $T_1=2$.

is the relaxation time of the magnetization m at temperature T_i plotted in Fig. 1(a). We find that the relative efficiencies of sampling in the actual ST and MC simulations is accurately predicted by Eq. (1) up to an upper temperature of $T_N=2.3$ (with $N=7$ temperatures, resulting in an efficiency gain of a factor of ≈ 30). For upper temperatures $T_N>2.3$, Eq. (1) slightly overestimates the efficiency gain. At these high temperatures, the magnetization is barely bimodal [Fig. 1(b), with a low barrier of only $\approx 1k_B T$ in the free energy profile along m], and both the assumption of two-state dynamics and fast temperature change begin to break down.

A comparison of the normalized autocorrelation function of the magnetization $C(t)$ from simulations and theory further demonstrates the quantitative accuracy of the kinetic model of ST (Fig. 4). We find that $C(t) = \langle \Delta m(t) \Delta m(0) \rangle / \langle \Delta m^2 \rangle$ with $\Delta m(t) = m(t) - \langle m \rangle$, as obtained from the ST simulations, matches the theoretical prediction of $C(t) \approx \exp[-(k_{\text{eff}}^+ + k_{\text{eff}}^-)t]$ with k_{eff}^\pm defined in Eq. (9). The correlation functions are calculated from the time series of $m(t)$ irrespective of the instantaneous temperatures in the ST simulations. Figure 4 shows that the magnetization relaxes exponentially with a rate of k_{eff} , with only small deviations even at a high upper temperature of $T_9=2.4$.

B. Comparison to replica exchange Monte Carlo

For comparison, we have also run REMC simulations with replicas at the same temperatures as in the ST simulations. Figure 3 shows that the resulting efficiency gains for sampling the magnetization at $T_1=2$ are nearly identical to those of the ST simulations, and again in full agreement with Eq. (1) up to an upper temperature of $T_N=2.3$.

C. Effect of tempering protocol

To test the influence of slowing down the sampling of the different temperatures T_i (i.e., the diffusion in temperature space), we have also performed ST simulations in which temperature changes are only allowed with neighboring temperatures ($j=i \pm 1$) instead of picking the trial temperature T_j randomly from the whole temperature range. For upper tem-

peratures $T_N < 2.3$, the efficiency gains are identical to those in ST simulations with random choice of the trial temperature T_j (not shown); for higher temperatures T_N , the efficiency gains fall below the theoretical limit of Eq. (1) and those in ST simulations with faster sampling of temperature space. We found that for $T_N = 2.45$, the relaxation time of the temperature in the ST simulations with neighbor switching is comparable to the relaxation time of the magnetization at T_N . As a consequence, the assumption of fast temperature change is no longer valid, explaining the reduced gain in efficiency. In essence, with temperature relaxation being slow, transitions between positive and negative magnetizations at a high temperature have a high probability of reverting before the ST system “cools down” to the temperature of interest.

D. Effect of the number of temperatures

A somewhat surprising prediction from the efficiency formula equation (1) is that for a fixed temperature range T_1 to T_N , and with a given distribution of temperatures (e.g., uniform), the efficiency gain η_m will approach a constant in the limit of large N .¹⁶ This means that even though the fraction Q_m of time spent at the target temperature decreases as $1/N$, the sampling efficiency η_m is predicted to be unaffected. However, Eq. (1) applies only in the limit of fast temperature change, which may not be the case for large N , depending on the tempering protocol used. If temperature changes are only performed between nearest neighbors in temperature, then for large N the mean first passage time (in units of attempted temperature changes) to diffuse from a high temperature T_N to a low temperature T_1 will scale as N^2 . Specifically, for uniform weights $Q_i = 1/N$, we can estimate the relaxation time of the temperature from the corresponding time for diffusion on a finite interval,

$$\tau_T \approx \frac{N^2 \Delta t}{10 \overline{p_{\text{acc}}}}, \quad (23)$$

where Δt is the time between attempted temperature changes in the same direction (e.g., up), and $\overline{p_{\text{acc}}}$ is the average probability of acceptance. Here we assume $\overline{p_{\text{acc}}}$ to be independent of temperature and the state of the system and $p(T_i|F)$ to be roughly constant. We found Eq. (23) to be an excellent approximation for the relaxation time of the temperature (Fig. 5).

For large N with only nearest-neighbor temperature exchanges the assumption of fast temperature change will be violated. We tested this prediction by performing simulations with a fixed temperature range, but covered with different numbers N of uniformly spaced temperatures. For $T_1 = 2$ and $T_N = 2.4$ with nearest-neighbor exchange, we found a significant gain in efficiency when using only $N = 3$ or 5 temperatures ($\eta_1 \approx 53$) instead of 9 ($\eta_1 \approx 30$), but still somewhat lower than what is obtained by picking the trial temperature T_j at random from $N = 9$ uniformly spaced temperatures ($\eta_1 \approx 59$). In contrast, global temperature changes, $i \rightarrow 1, 2, \dots, i-1, i+1, \dots, N$, result in much faster temperature relaxation (Fig. 5). In practice, one should monitor and

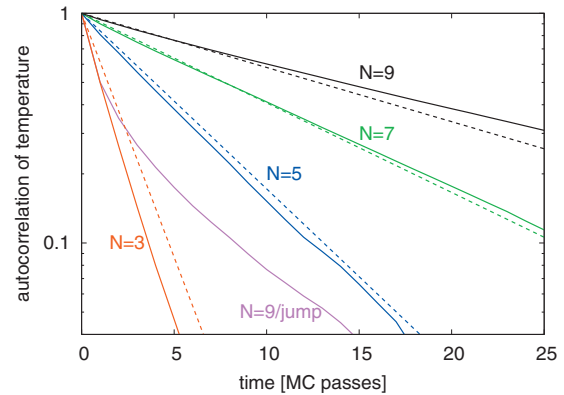


FIG. 5. Normalized autocorrelation function of the temperature in ST simulations with nearest-neighbor temperature exchanges (solid lines), $N = 3, 5, 7$, and 9 evenly spaced temperatures and upper temperatures of $T_N = 2.1, 2.2, 2.3$, and 2.4. Dashed lines show the theoretical predictions $\exp(-t/\tau_T)$ with τ_T from Eq. (23). Also shown is the curve for $N = 9$ with global temperature changes ($i \rightarrow 1, 2, \dots, i-1, i+1, \dots, N$) permitting distant jumps in the temperature.

optimize the rate of temperature relaxation [e.g., from the autocorrelation function of $T(t)$] and compare it to the rate of relaxation of the two-state process. Fast temperature change is essential to achieve the maximum possible efficiency.

E. Effect of temperature weights

According to Eq. (1), the computational efficiency can be increased by giving higher weights Q_i to temperatures T_i , where $\tau_i^+ + \tau_i^-$ is small and the system crosses frequently between the states. Typically, these will be the higher temperatures. We have tested Eq. (1) for nonuniform Q_i by running ST simulations of the Ising system with linearly increasing weights, $Q_i = i / \sum_{j=1}^N j$. We again find excellent agreement between the predicted and observed gains in efficiency up to about $N = 7$ with an upper temperature of $T_7 = 2.3$ (data not shown). For higher upper temperatures, the gains are somewhat below those predicted, as with uniform weights. Overall, using a linear bias toward higher temperatures results in gains $\approx 35\%$ higher than those for uniform weights. Gains from increasing the weight of the higher temperatures were previously observed in simulations by Zhang and Ma.²⁶

F. Accelerated tempering protocol

We have also tested a scheme to speed up the temperature relaxation. Instead of first picking the temperatures T_j from $1, 2, \dots, i-1, i+1, \dots, N$ (or from $j = \pm 1$) and then accepting a change from T_i to T_j according to Eq. (1), we choose the temperatures T_j at random according to their probability for a given conformation with energy E ,

$$p(T_j|E) = \frac{\exp(-E/k_B T_j + f_j)}{\sum_{k=1}^N \exp(-E/k_B T_k + f_k)}. \quad (24)$$

To sample T_j from this distribution, we calculate the cumulative distribution $P_j = \sum_{k=1}^j p(T_k|N)$, with $P_0 = 0$ and $P_N = 1$. We then draw a random number r with uniform distribution between 0 and 1. The new temperature T_j is determined such

that $P_{j-1} < r \leq P_j$. Care has to be taken to avoid over-/underflows in the numerical calculation, e.g., by adding a suitable constant to all energies E . As an alternative, one can also repeatedly apply the standard ST temperature change protocol for a fixed configuration of energy E , and with detailed balance properly enforced.

In simulations of the Ising model with $N=10$ temperatures, and $T_1=2$ and $T_{10}=2.45$, the efficiency gains from this protocol over picking trial temperatures at random are negligible [$\eta_1 \approx 58$ in both cases, compared to $\eta_1=88$ from Eq. (1)]. We conclude from this result (and similar results from simulations run with attempted temperature changes after every attempted spin flip) that the gradual breakdown of the two-state assumption is primarily responsible for the actual efficiency being slightly lower than predicted for upper temperatures $T_N > 2.3$.

IV. CONCLUDING REMARKS

We have derived an analytical expression for the relative computational efficiency of ST simulations over regular MC or MD simulations. The expression, Eq. (1), applies to the important case that the slow dynamics of the system can be described by the relaxation between two metastable states, and is derived under the assumption that the temperature change in ST is fast compared to the transitions between the two states. We also sketch an extension to systems requiring more than two states. In our efficiency calculation, we have not considered the possibility of reweighting the data collected at the different temperatures. Combining data from all temperatures, for instance, by using the weighted histogram method,^{27,28} can further enhance the sampling efficiency at the temperature of interest, in particular if ST is not in the fast limit; but gains should be minimal if ST is fast with respect to the property of interest.

We have tested the expression for the efficiency gain against MC calculations of the magnetization in a 2D Ising model. We have found that Eq. (1) predicts the efficiency gains quantitatively. Only at the highest upper temperatures T_N considered, the observed sampling efficiency gains are somewhat below those predicted. The reasons for this drop are, on the one hand, that the assumption of fast temperature change is increasingly violated at high T_N (where the relaxation between up and down magnetizations becomes faster than the relaxation from high to low temperatures); and, on the other hand, that the two-state assumption for the magnetization breaks down as the temperature T_N exceeds T_c .

Interestingly, the formula equation (1) for the efficiency gain in ST simulations is essentially identical to that derived previously on a very different route for REMD and REMC simulations. However, in the limit of fast exchange and a large number N of replicas, we have shown that the equivalence of the two approaches follows from simple statistical arguments. As an interesting consequence, in the asymptotic limit running multiple ST simulations in parallel without communication is as efficient as running REMD or REMC simulations with communication for replica exchange. This makes ST particularly useful for distributed computing.¹⁵

Our analytical result for the efficiency gain in ST simulations has practical implications in the setup of ST simulations as follows.

A. Temperature range T_1 to T_N

As in REMC and REMD simulations,¹⁶ one should cover a range of temperatures T_i , where the sum of relaxation times $\tau_i^+ + \tau_i^-$ is smaller than $\tau_m^+ + \tau_m^-$ at the temperature of interest, T_m . In this way, the more frequent transitions (corresponding to a higher reactive flux) at the higher temperatures result in improved sampling at the lower temperatures of interest. To further maximize η_m with respect to T_1 and T_N , the strategies discussed in Sec. II.E of our previous work¹⁶ for REMD and REMC can be adapted.

Importantly, because the sum of lifetimes $\tau_i^+ + \tau_i^-$ enters the efficiency formula equation (1), not the relaxation rate $k_i^+ + k_i^-$, ST, REMC, and REMD are not as powerful to enhance the sampling of entropy dominated processes such as protein folding as they are for enthalpy dominated transitions. Over a fairly broad range of temperatures, the folding rate of proteins is often found to be nearly constant or even slowing down with increasing temperature. Despite the acceleration of the enthalpy dominated unfolding process at high temperatures, the efficiency gain is limited by the unaccelerated folding process, resulting in long lifetimes of the unfolded state.

B. Temperature weights

Traditionally, one attempts to find weight factors f_i to ensure equal sampling $Q_i \approx 1/N$ for all temperatures. From the perspective of sampling efficiency, equal weighting is not optimal. Based on Eq. (1), it is advantageous to give higher weights to temperatures T_i at which $\tau_i^+ + \tau_i^-$ is smaller than $\tau_m^+ + \tau_m^-$ at the temperature of interest, T_m . In typical problems, one would thus want to give a disproportionately high weight to the higher temperatures. However, any optimization of the target weights Q_i has to consider that the relaxation in temperature space should be fast, as discussed next.

C. Temperature relaxation

In addition to choosing the optimal temperature range, it is important to ensure that temperature relaxation is fast compared to the relaxation of the system state at any given temperature [as has been found before for REMD (Refs. 16 and 19–21)]. Temperature relaxation can be accelerated by spacing the temperatures T_i more narrowly (if the acceptance of temperature changes is low), by making more frequent temperature change attempts, by using nonlocal temperature change moves (that must, however, satisfy detailed balance, e.g., by picking the trial temperatures T_j at random), or by picking temperatures T_j according to Eq. (24). Alternative implementations of ST,^{14,29,30} often related to approaches in REMD and REMC simulations,³¹ may also help enhance the sampling of temperature space, and in turn configuration space.

There has been some discussion in the literature of the optimal frequency of attempted temperature exchanges, in particular, in the context of replica exchange

simulations.^{21,32,33} On formal grounds, and consistent with other studies,^{21,26,33} we conclude that it is optimal to attempt temperature changes in ST, and temperature exchanges in REMC and REMD, as frequent as possible (as long as the computational cost of the exchange itself can be ignored). However, in practical implementations it is important that exchanges maintain a proper Boltzmann distribution. This is a particular concern in MD simulations, where one should use proper thermostats¹⁰ and integrators that evaluate positions and momenta at the same time step, such as velocity Verlet,³⁴ for proper rescaling (or redrawing) of the momenta.

D. Number N of temperatures

Interestingly, for a large number of temperatures N spaced according to a given distribution in a fixed range T_1 to T_N , the efficiency gain becomes independent of N , as long as the temperature change is fast. As a consequence, there is no loss in sampling efficiency from increasing the number N of temperatures, as long as the temperature relaxation rate does not suffer. However, circumventing the N^2 scaling of the temperature relaxation time for large N will require non-local temperature changes [such as the use of Eq. (24)].

In summary, we have shown that ST can substantially enhance the sampling efficiency over regular MC and MD simulations, as measured by a decrease in the simulation time required to achieve a certain statistical accuracy. In the limit of fast temperature change, this gain is largest, and equivalent to that in REMD and REMC simulations covering the same temperatures. Achieving high computational efficiency in ST simulations requires a careful selection of the simulation parameters, a process that can be guided by the analytical expression Eq. (1) for the maximum gain achievable.

ACKNOWLEDGMENTS

We thank Dr. Attila Szabo for many helpful and stimulating discussions and Dr. William Swope for discussions about the effects of integrators. This research was supported by the Intramural Research Program of the NIDDK, NIH.

- ¹A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, *J. Chem. Phys.* **96**, 1776 (1992).
- ²E. Marinari and G. Parisi, *Europhys. Lett.* **19**, 451 (1992).
- ³C. J. Geyer, *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface* (American Statistical Association, New York, 1991), pp. 156–163.
- ⁴K. Hukushima and K. Nemoto, *J. Phys. Soc. Jpn.* **65**, 1604 (1996).
- ⁵Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **314**, 141 (1999).
- ⁶A. E. García and K. Y. Sanbonmatsu, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2782 (2002).
- ⁷H. Fukunishi, O. Watanabe, and S. Takada, *J. Chem. Phys.* **116**, 9058 (2002).
- ⁸C. H. Bennett, *J. Comput. Phys.* **22**, 245 (1976).
- ⁹R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.* **57**, 2607 (1986).
- ¹⁰E. Rosta, N.-V. Buchete, and G. Hummer, *J. Chem. Theory Comput.* **5**, 1393 (2009).
- ¹¹U. H. E. Hansmann and Y. Okamoto, *Phys. Rev. E* **54**, 5863 (1996).
- ¹²U. H. E. Hansmann and Y. Okamoto, *J. Comput. Chem.* **18**, 920 (1997).
- ¹³A. Mitsutake, Y. Sugita, and Y. Okamoto, *Biopolymers* **60**, 96 (2001).
- ¹⁴S. Rauscher, C. Neale, and R. Pomés, *J. Chem. Theory Comput.* **5**, 2640 (2009).
- ¹⁵X. H. Huang, G. R. Bowman, and V. S. Pande, *J. Chem. Phys.* **128**, 205106 (2008).
- ¹⁶E. Rosta and G. Hummer, *J. Chem. Phys.* **131**, 165102 (2009).
- ¹⁷S. Park and V. S. Pande, *Phys. Rev. E* **76**, 016703 (2007).
- ¹⁸W. Zheng, M. Andrec, E. Gallicchio, and R. M. Levy, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15340 (2007).
- ¹⁹W. Nadler and U. H. E. Hansmann, *J. Phys. Chem. B* **112**, 10386 (2008).
- ²⁰W. Nadler and U. H. E. Hansmann, *Phys. Rev. E* **75**, 026109 (2007).
- ²¹D. Sindhikara, Y. L. Meng, and A. E. Roitberg, *J. Chem. Phys.* **128**, 024103 (2008).
- ²²M. Boguñá, A. M. Berezhkovskii, and G. H. Weiss, *J. Phys. Chem. A* **105**, 4898 (2001).
- ²³A. M. Berezhkovskii, A. Szabo, and G. H. Weiss, *J. Chem. Phys.* **110**, 9145 (1999).
- ²⁴I. V. Gopich and A. Szabo, *J. Chem. Phys.* **118**, 454 (2003).
- ²⁵H. Nymeyer, *J. Chem. Theory Comput.* **4**, 626 (2008).
- ²⁶C. Zhang and J. P. Ma, *J. Chem. Phys.* **129**, 134112 (2008).
- ²⁷A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.* **63**, 1195 (1989).
- ²⁸J. D. Chodera, W. C. Swope, J. W. Pitera, C. Seok, and K. A. Dill, *J. Chem. Theory Comput.* **3**, 26 (2007).
- ²⁹G. Stolovitzky and B. J. Berne, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11164 (2000).
- ³⁰R. Denschlag, M. Lingenheil, P. Tavan, and G. Mathias, *J. Chem. Theory Comput.* **5**, 2847 (2009).
- ³¹P. Liu, B. Kim, R. A. Friesner, and B. J. Berne, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13749 (2005).
- ³²M. J. Abraham and J. E. Gready, *J. Chem. Theory Comput.* **4**, 1119 (2008).
- ³³X. Periole and A. E. Mark, *J. Chem. Phys.* **126**, 014903 (2007).
- ³⁴W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, *J. Chem. Phys.* **76**, 637 (1982).