

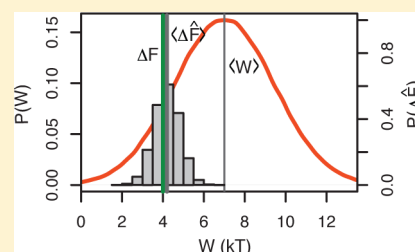
Estimation of Free-Energy Differences from Computed Work Distributions: An Application of Jarzynski's Equality

Ignacia Echeverria[†] and L. Mario Amzel*

Department of Biophysics and Biophysical Chemistry, School of Medicine, Johns Hopkins University, Baltimore, Maryland 21205, United States

Supporting Information

ABSTRACT: Equilibrium free-energy differences can be computed from non-equilibrium molecular dynamics (MD) simulations using Jarzynski's equality (Jarzynski, *C. Phys. Rev. Lett.* **1997**, *78*, 2690) by combining a large set of independent trajectories (path ensemble). Here we present the multistep trajectory combination (MSTC) method to compute free-energy differences, which by combining trajectories significantly reduces the number of trajectories necessary to generate a representative path ensemble. This method generates well-sampled work distributions, even for large systems, by combining parts of a relatively small number of trajectories carried out in steps. To assess the efficiency of the MSTC method, we derived analytical expressions and used them to compute the bias and the variance of the free-energy estimates along with numerically calculated values. We show that the MSTC method significantly reduces both the bias and variance of the free-energy estimates compared to the estimates obtained using single-step trajectories. In addition, because in the MSTC method the process is divided into steps, it is feasible to compute the reverse transition. By combining the forward and reverse processes, the free-energy difference can be computed using the Crooks' fluctuation theorem (Crooks, G. E. *J. Stat. Phys.* **1998**, *90*, 1481 and Crooks, G. E. *Phys. Rev. E* **2000**, *61*, 2361) or Bennett's acceptance ratio (Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245), which further reduces the bias and variance of the estimates.



INTRODUCTION

An accurate estimation of free-energy differences (ΔF) by computational methods is central to understanding the details of the interactions contributing to the thermodynamic and statistical mechanical descriptions of processes involving molecular systems ranging from liquids to macromolecules. In particular, molecular dynamics (MD) simulations have proven to be a powerful tool for estimating ΔF values between two states and shedding light on the factors and interactions that define these states.^{1–11} Given two states A and B, the free-energy difference $\Delta F = F_B - F_A$ determines the relative probability of the system adopting one state or the other. The free-energy difference between these states can be calculated as

$$\Delta F = F_B - F_A = -\beta^{-1} \ln \left(\frac{Q_B}{Q_A} \right) \quad (1)$$

where $\beta = 1/(kT)$ (k is the Boltzmann constant and T is the temperature) and Q_A and Q_B are the partition functions of states A and B. In a broad sense, path-based methods used to calculate ΔF can be divided into two categories according to the way that the path is sampled: (1) equilibrium sampling of states A and B or (2) driving the system (reversibly or irreversibly) between states A and B. The first approach is based on the fact that the ratio of partition functions (eq 1) is proportional to the ratio of probabilities of occurrence of states A and B. Examples are the free-energy perturbation (FEP) method^{12–15} and methods based on the potential of mean force

(PMF) combined with umbrella sampling (US).¹⁶ In all of these methods, without a loss of generality, the transition between states A and B can be further subdivided into “windows” and the probability of occurrence of different states is calculated for each of them. Examples of the second approach are thermodynamic integration (TI)^{17–19} (where the transition is done in such a way that a series of equilibrium configurations is sampled between states A and B) and methods based on nonequilibrium perturbations.^{20–24} When the methods used to calculate ΔF values are based on sampling a specific path that connects states A and B, the calculations provide mechanistic details of the transition and the associated free-energy surface. Therefore, free-energy calculations using MD simulations make it possible to understand the underlying processes at the atomic level and probe states of a system that are not accessible experimentally. (For a review of free-energy calculations, see the book by Chipot and Pohorille.²⁵)

When a system is driven along a path connecting two states, the transition can be characterized by the amount of work needed to drive the transition. In general, the average work satisfies the relation $\langle W \rangle \geq \Delta F$, where the equality holds for reversible processes. In 1997, Jarzynski^{26,27} showed that even for irreversible processes ΔF can be calculated from the work

Received: January 16, 2012

Revised: July 27, 2012

Published: July 31, 2012

done to drive the system between two equilibrium states A and B using the relation

$$\Delta F = -\beta^{-1} \ln \langle e^{-\beta W} \rangle \quad (2)$$

where $\langle \cdot \rangle$ denotes an ensemble average over a large number of realizations that drive the system from equilibrium state A to equilibrium state B.^{26,27} The system may be driven between two states by, for example, a time-dependent Hamiltonian. This equation sets no constraints on how the system is driven between states A and B and, remarkably, can be used to obtain equilibrium free-energy differences using nonequilibrium perturbations as long as the initial and final states are at equilibrium. Note that Jarzynski's equation does not necessarily require the system to reach an equilibrium final state but indicates that the control parameter attains its final value, whereas the system may relax to equilibrate to a well-defined state that is consistent with the final value of the control parameter. Nevertheless, in the thermodynamic sense free-energy differences are defined only between equilibrium states. Jarzynski's equality (JE) is an exact result and has been proven for various thermodynamic ensembles and under different sets of assumptions^{28–30} and tested experimentally.^{31–34} For a stochastic process, a different work value will be associated with each realization. In practical terms, N repetitions of a thermodynamic process can be obtained from, for example, single-molecule experiments in which repeated nonequilibrium force measurements are made or by molecular dynamics simulations where, starting from an initial equilibrium state A sampled from a Boltzmann distribution the system is driven to a final state B by “switching” parameter λ at a finite rate. For a data set that contains N statistically independent measurements of the work needed to drive the transition ($\{W_1, W_2, \dots, W_N\}$), the free-energy estimate can be approximated by

$$\Delta \hat{F}(N) \approx -\beta^{-1} \ln \frac{1}{N} \sum_i e^{-\beta W_i} \quad (3)$$

where the sum is done over N independent work values. In this form, JE constitutes a powerful estimator of the free-energy difference from a set of N irreversible transitions between two states A and B, and no information is needed regarding the distribution of work values. This approximation becomes an equality when $N \rightarrow \infty$.

The set of all possible realizations provides a work probability distribution ($P(W)$) that can be used to calculate the free-energy difference as

$$\Delta F = -\beta^{-1} \ln \int dW e^{-\beta W} P(W) \quad (4)$$

The Jarzynski exponential average heavily weights the tail of the work probability distribution (low W values), which is in general not well sampled. This is illustrated in Figure 1 (top panel).

A comparison of eqs 3 and 4 shows that JE is valid, provided that enough repetitions of the process can be carried out.^{35–38} Unfortunately, the number of repetitions required to obtain reliable results can be too large to be realistic to achieve, especially in situations involving large fluctuations. Several studies have evaluated the convergence of free-energy estimates^{13,19,39–42} by quantifying the bias (the difference between the estimated expected value and the true value, i.e., the accuracy) and the variance (a measure of the spread of the individual estimated values, i.e., the precision) of the estimates

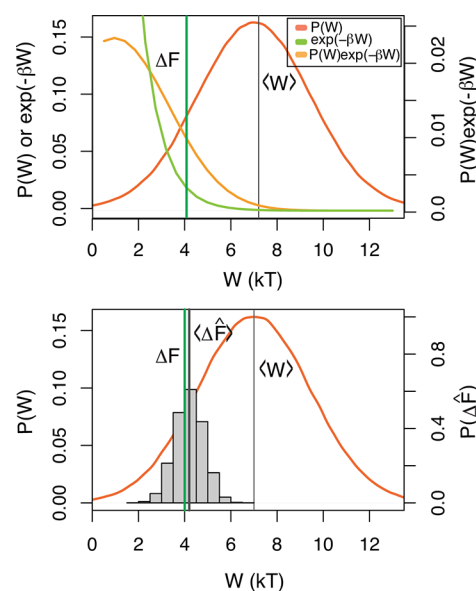


Figure 1. (Top) Jarzynski average of the exponential work values. Both the work probability distribution (red curve) and the exponential of the work values $\exp(-\beta W)$ (green curve) are shown. Jarzynski's equality heavily weights the low tail of the work distribution as shown by the product $P(W) \exp(-\beta W)$ (orange curve). Vertical lines represent the average work (solid line) and the estimated free energy (green line). (Bottom) Distribution of the free-energy estimates. The work probability distribution is shown in red; the values of $\Delta \hat{F}$ in different simulations are shown as a histogram. Vertical lines represent the average work (thin line), the true free energy (green line), and the estimated free energy (thick solid line).

when the free energy is extracted from data obtained experimentally or by simulation.

In this article, we present and evaluate the accuracy of the free-energy differences estimated by the multistep trajectory combination (MSTC) method using JE.^{26,27} Previous studies have shown that a well-sampled work distribution can be obtained by dividing long trajectories into shorter steps and equilibrating at the end of each step.^{19,43,44} These well-sampled work distributions have later been used to obtain accurate free-energy differences using Jarzynski's equation. Despite this early success, the estimation of the number of steps and the number of trajectories required in any particular case remains completely empirical. More to the point, the relation between these two parameters and the expected error in the estimation of the free-energy change has not been derived. The general scheme consists of dividing the trajectory between the initial and final states into M steps and equilibrating the system at the end of each step. After generating N trajectories of M steps each, new paths are generated by combining steps of different trajectories, taking advantage of the fact that because the system has been equilibrated with identical external conditions, microstates of different trajectories at the end of each step belong to the same macrostate. During these equilibrations, the average properties of the system (macrostate) fluctuate about the same values in the different trajectories. Thus, even if the time is not long enough for the system to sample some of the same conformations in all trajectories, the states accessed in the different trajectories do belong to the same ensemble. This method can be used to obtain well-sampled work distributions, even for large systems, from a relatively small number of trajectories carried out in steps. Here we show that the MSTC

method significantly reduces the bias and the variance of the free-energy estimates. In other words, given a certain bias or variance threshold, far fewer trajectories are needed to achieve comparable results when the MSTC method is used. In addition, the fact that trajectories are divided into steps also makes it possible to compute the reverse transition when using MD simulations by reversing each of the steps in the forward process. When either Crooks' fluctuation theorem (CFT)²⁸ or Bennett's acceptance ratio (BAR)⁴⁵ is used with the forward and reverse transitions, the bias and the variance of the free-energy estimates are further reduced. The results presented in this work form the basis for the application of the MSTC method to a variety of ΔF calculations using MD simulations. Relating the work fluctuations to the bias and variance of the estimates provides the tools to determine the number of trajectories needed for an accurate estimation of free-energy differences.

THEORY AND METHODS

MSTC Method: Thermodynamic Background. The application of the MSTC method requires the generation of a set of N trajectories of M steps each. These trajectories can be combined to produce N^M trajectories, which are used to compose a work probability distribution ($P(W)$). The computation of $\Delta\hat{F}$ over the N^M trajectories is equivalent to computing $\Delta\hat{F}_M$, the value of the free energy of each of the M steps averaged over the N trajectories, and adding them to obtain the total $\Delta\hat{F} = \sum_{i=1}^M \Delta\hat{F}_i$.

This can be proven by considering a set of N trajectories of M steps each. The work done during one step is given by $W_{s,t}$ with s being the step index and t being the trajectory index. Using this nomenclature, Jarzynski's equality can be rewritten by summing over all N^M trajectories

$$e^{-\beta\Delta F} = \langle e^{-\beta W} \rangle \quad (5)$$

$$= \left(\frac{1}{N} \sum_{t=1}^N e^{-\beta W_{1,t}} \right) \left(\frac{1}{N} \sum_{t=1}^N e^{-\beta W_{2,t}} \right) \dots \left(\frac{1}{N} \sum_{t=1}^N e^{-\beta W_{M,t}} \right) \quad (6)$$

$$= \frac{1}{N^M} \prod_{s=1}^M \sum_{t=1}^N e^{-\beta W_{s,t}} \quad (7)$$

where steps are considered to be statistically independent and therefore the covariance between steps is assumed to be 0. Then, the free-energy difference for a set of N trajectories of M steps each is given by

$$\Delta F_{\text{MSTC}} = \sum_{s=1}^M -\beta^{-1} \ln \frac{1}{N} \sum_{t=1}^N e^{-\beta W_{s,t}} \quad (8)$$

Hereafter, we refer to this as the free-energy estimation from the multistep trajectory combination (MSTC). This result can be compared to the free-energy estimate obtained when only the original set of N trajectories is considered, without using any trajectory combinations. In that case, the work for each trajectory is just the sum over steps

$$W_t = \sum_{s=1}^M W_{s,t} \quad (9)$$

and Jarzynski's equality can be written as

$$\Delta F_{\text{OSP}} = -\beta^{-1} \ln \frac{1}{N} \sum_{t=1}^N e^{-\beta \sum_{s=1}^M W_{s,t}} \quad (10)$$

Hereafter, we refer to this as the free-energy estimation from one-step processes (OSP).

Free-energy estimates based on exponential averaging are always biased when only a finite number of trajectories are available. Several studies have quantified different measures of systematic error such as the bias, the variance, and the mean-squared error.^{13,39,40,42} To assess the bias and variance of the free-energy estimates given by eqs 8 and 10, we followed the approach presented by Gore et al.³⁹ in which different approximations are used in the small- N and large- N regimes. This is particularly relevant because whereas in the large- N regime neither the bias nor the variance is a significant source of error³⁹ the N required for this to be true is usually inaccessible by experimental or computational methods. Therefore, achieving a low bias and variance in the small- N regime is crucial for accurate estimations of free-energy differences.

To calculate the expected value of the free-energy estimator ($\langle \Delta\hat{F} \rangle$), we defined the variable $X = \langle e^{-\beta W} \rangle_N$ and assumed that its variance $\text{Var}(e^{-\beta W})$ is finite. Also, for a sufficiently large N , the central limit theorem guarantees that X will be normally distributed with mean $\bar{X} = e^{-\beta\Delta F}$ and variance $\text{Var}(X) = \text{Var}(e^{-\beta W})/N$. If we expand $\ln(X)$ around \bar{X} ,

$$\Delta F_{\text{JE}} = \beta^{-1} \left(\ln(\bar{X}) + \frac{1}{\bar{X}}(X - \bar{X}) - \frac{1}{2\bar{X}^2}(X - \bar{X})^2 + \dots \right) \quad (11)$$

then the expected value of the free-energy estimate up to second order is given by

$$\langle \Delta\hat{F}_{\text{JE}} \rangle = \Delta F_{\text{JE}} + \frac{\text{Var}(e^{-\beta W})}{2\beta N \langle e^{-\beta W} \rangle^2} \quad (12)$$

Free-energy estimators (eqs 10 and 8) can be rewritten by considering that $\text{Var}(X) = E[X^2] - (E[X])^2$, where $E[\cdot]$ represents the expected value. Thus, for a set of N trajectories of M steps,

$$\langle \Delta\hat{F}_{\text{OSP}} \rangle = \sum_{s=1}^M \Delta F_s + \frac{\prod_{s=1}^M E[e^{-2\beta W_s}] - \prod_{s=1}^M (E[e^{-\beta W_s}])^2}{2\beta N \prod_{s=1}^M (E[e^{-\beta W_s}])^2} \quad (13)$$

and

$$\langle \Delta\hat{F}_{\text{MSTC}} \rangle = \sum_{s=1}^M \Delta F_s + \sum_{s=1}^M \frac{E[e^{-2\beta W_s}] - (E[e^{-\beta W_s}])^2}{2\beta N (E[e^{-\beta W_s}])^2} \quad (14)$$

Significantly, in eqs 13 and 14 the first term corresponds to ΔF . Then, given that the bias is defined as $B(N) = \langle \Delta\hat{F}(N) \rangle - \Delta F$, the second term corresponds to the bias.

For a known work probability distribution $P(W)$, it is possible to calculate the expectation values from eqs 13 and 14 analytically. For example, in the near-equilibrium regime the work probability distribution $P(W)$ will be normally distributed. This general property is valid for perturbations performed sufficiently slowly such that the perturbation is reversible. In this case, the work values for every step are random variables $W_{s,t} \sim \mathcal{N}(\mu_s, \sigma_s^2)$, where μ_s is the mean and σ_s^2 is the variance of the normal distribution \mathcal{N} for a given step s . If we consider that all steps are independent random variables, then the sum

of these random variables is also normally distributed. Hence, the total work done in one trajectory of M steps is given by $W_t = \sum_{s=1}^M W_{s,t} \sim \mathcal{N}(\sum_s^M \mu_s, \sum_s^M \sigma_s^2)$. (For clarity, we will use σ^2 for the variance of a distribution and $\text{Var}(\cdot)$ for the variance of an estimate.) Considering these assumptions, we can rewrite eqs 13 and 14 as

$$\langle \Delta \hat{F}_{\text{OSP}} \rangle = \sum_{s=1}^M \Delta F_s + \frac{e^{\beta^2 \sigma_T^2} - 1}{2\beta N} \quad (15)$$

and

$$\langle \Delta \hat{F}_{\text{MSTC}} \rangle = \sum_{s=1}^M \Delta F_s + M \frac{e^{\beta^2 \sigma_T^2 / M} - 1}{2\beta N} \quad (16)$$

where $\sigma_T = \sum_s^M \sigma_s^2$.

As already mentioned, a different approach is necessary in the small- N regime. To determine $\langle \Delta \hat{F} \rangle$, we considered the case where only one trajectory is available. The bias will be given by $B(N=1) = \langle W \rangle - \Delta F = \bar{W}_{\text{dis}} = 1/2\beta\sigma_T^2$, where \bar{W}_{dis} is the average dissipated work. Then, in the small- N regime a reasonable approximation of the bias is

$$B(N) \approx \frac{\bar{W}_{\text{dis}}}{N^\alpha} \quad (17)$$

where α is function of \bar{W}_{dis} ($\alpha = \alpha(\bar{W}_{\text{dis}})$). To determine the constant α , boundary conditions between the small- N and large- N regimes were considered. Following the procedure of Gore et al.,³⁹ we considered the near-equilibrium results from eqs 15 and 16. If we first consider the OSP process, it is expected that in the large- N regime $N \gg (e^{2\beta\bar{W}_{\text{dis}}} - 1)$ and that the two regimes will intersect at $N_{\text{int}} = C_{\text{OSP}}(e^{2\beta\bar{W}_{\text{dis}}} - 1)$. Then combining the expressions for the bias in the large- N (eq 15) and small- N (eq 17) regimes at N_{int} yields

$$B_{\text{OSP}}(N_{\text{int}}) = \frac{e^{2\beta\bar{W}_{\text{dis}}} - 1}{2\beta N_{\text{int}}} = \frac{\bar{W}_{\text{dis}}}{N_{\text{int}}^{\alpha_{\text{OSP}}}} = \frac{1}{2\beta C_{\text{OSP}}} \quad (18)$$

By solving this equation for α_{OSP} we obtain

$$\alpha_{\text{OSP}} = \frac{\ln(2\beta C_{\text{OSP}} \bar{W}_{\text{dis}})}{\ln(C_{\text{OSP}}(e^{2\beta\bar{W}_{\text{dis}}} - 1))} \quad (19)$$

where the single constant C_{OSP} will determine where the two regimes intersect. C_{OSP} is the only free parameter of the model and was determined empirically. The constant α_{OSP} is interpreted as the rate at which the bias falls to zero.

Analogously, for the MSTC process in the large- N regime $N \gg (e^{2\beta\bar{W}_{\text{dis}}/M} - 1)$, and the two regimes will intersect at $N_{\text{int}} = C_{\text{MSTC}}(e^{2\beta\bar{W}_{\text{dis}}/M} - 1)$. By applying the boundary condition, we obtain

$$B_{\text{MSTC}}(N_{\text{int}}) = M \frac{e^{2\beta\bar{W}_{\text{dis}}/M} - 1}{2\beta N_{\text{int}}} = \frac{\bar{W}_{\text{dis}}}{N_{\text{int}}^{\alpha_{\text{MSTC}}}} = \frac{M}{2\beta C_{\text{MSTC}}} \quad (20)$$

Again, by solving for α_{MSTC} we obtain

$$\alpha_{\text{MSTC}} = \frac{\ln(2\beta C_{\text{MSTC}} \bar{W}_{\text{dis}}/M)}{\ln(C_{\text{MSTC}}(e^{2\beta\bar{W}_{\text{dis}}/M} - 1))} \quad (21)$$

Comparison of equations 19 and 21 show that α_{MSTC} is always bigger than α_{OSP} .

In summary, in the near-equilibrium regime, the free-energy bias including terms to second order is given by

$$B_{\text{OSP}}(N) = \begin{cases} \frac{\bar{W}_{\text{dis}}}{N^{\alpha_{\text{OSP}}}}, & \text{for } 0 < N < C_{\text{OSP}}(e^{2\beta\bar{W}_{\text{dis}}} - 1) \\ \frac{e^{2\beta\bar{W}_{\text{dis}}} - 1}{2\beta N}, & \text{for } C_{\text{OSP}}(e^{2\beta\bar{W}_{\text{dis}}} - 1) < N < \infty \end{cases} \quad (22)$$

and

$$B_{\text{MSTC}}(N) = \begin{cases} \frac{\bar{W}_{\text{dis}}}{N^{\alpha_{\text{OSP}}}}, & \text{for } 0 < N < C_{\text{MSTC}}(e^{2\beta\bar{W}_{\text{dis}}/M} - 1) \\ M \frac{e^{2\beta\bar{W}_{\text{dis}}/M} - 1}{2\beta N}, & \text{for } C_{\text{MSTC}}(e^{2\beta\bar{W}_{\text{dis}}/M} - 1) < N < \infty \end{cases} \quad (23)$$

To estimate the variance of the free energy, it is necessary to consider again both the large- and small- N regimes. In the large- N regime, if we substitute $\sigma^2(N) = \langle (\Delta \hat{F} - \langle \Delta \hat{F} \rangle)^2 \rangle$ into eq 11 we obtain

$$\text{Var}_{\text{OSP,MSTC}}(N) = \frac{2B_{\text{OSP,MSTC}}(N)}{\beta} \quad (24)$$

In the small- N regime, if we again consider the case of the $N = 1$ distribution, then $\text{Var}(N = 1) = \sigma_W^2 = 2\bar{W}_{\text{dis}}/N = 2B(N)/\beta$. Following an equivalent approach to that presented for estimating the bias, we obtain

$$\text{Var}_{\text{OSP,MSTC}}(N) = \frac{2\bar{W}_{\text{dis}}}{\beta N^{\alpha_{\nu, \text{OSP,MSTC}}}} \quad (25)$$

where α_{ν} has the same form as α (eq 19 or 21), but the constant C_{ν} has a different value.

This procedure can be extended to other work probability distributions. The analytical expressions for the bias and variance for a gamma distribution have been included in the Supporting Information.

Numerical Calculations: Bias and Variance Estimations. By using numerically generated work distributions, we computed both the free-energy bias ($B(N)$) and the variance ($\text{Var}(N)$) for different distributions and compared them to the results derived in the previous sections. This is illustrated in Figure 1 (lower panel), where the work probability distribution and the histogram of the $\Delta \hat{F}$ values obtained numerically are shown. In the figure, the difference between ΔF and $\Delta \hat{F}$ corresponds to the bias of the free-energy estimates and the spread of the estimates obtained represents the variance.

To estimate the bias and variance, an $N \times M$ matrix of work values was generated for a set of N trajectories of M steps. The work values associated with every step were randomly sampled from a given distribution, and without a loss of generality, the same distribution was assumed for every step.

The bias ($B(N)$) and variance ($\text{Var}(N)$) of the free-energy estimates were computed for normal distributions of various total variances ($\sigma_T^2 = (kT)^2, 2(kT)^2, 4(kT)^2, 8(kT)^2, 16(kT)^2$, and $32(kT)^2$). The number of trajectories considered in our calculations ranged from 1 to 10^5 . The number of steps was $M = 10$. Each ΔF calculations was repeated between 1000 and 5000 times until the bias and variance had converged. In addition, we compared the results in cases where all steps were described by the same distribution (same variance) and cases where one step was sampled from distributions with a higher

variance but by conserving the total variance in the work values for each trajectory. In addition, the bias and variance in the free-energy estimation were computed for gamma distributions.

Free-Energy Estimations by Combining Forward and Reverse Simulations. Crooks' fluctuations theorem (CFT)²⁸ is a general nonequilibrium relation that can be considered to be a generalization of Jarzynski's equality. This theorem states that

$$\frac{P_F(W)}{P_R(-W)} = e^{\beta(W-\Delta F)} \quad (26)$$

where $P_F(W)$ denotes the work probability distribution for a forward process, $P_R(-W)$ denotes that for the reverse process, and ΔF is the free-energy change between the final and initial states and is equal to the reversible work associated with the process. This symmetry relation in the work fluctuations between the forward and reverse processes does not set any constraints on how far from equilibrium the system is driven, as long as the process starts and ends in an equilibrium state. It is clear from eq 26 that ΔF can be obtained directly from the intersection of the forward and reverse distributions.

Also, Bennett⁴⁵ demonstrated that when the forward and reverse work probability distributions are available it is possible to use these two distributions to obtain the free energy of the process. Using Bennett's protocol, one obtains estimates that are significantly better than those obtained when only the forward work distribution is used. Furthermore, Pande and co-workers⁴¹ showed that the Bennett acceptance ratio (BAR) method can be interpreted as the maximum likelihood estimate of the free-energy difference, given the forward and reverse work probability distributions. The maximum likelihood estimate is obtained when

$$\frac{\partial \ln L(\Delta F)}{\partial \Delta F} = \sum_{i=1}^{N_F} \frac{1}{1 + \exp(\beta(M + W_i - \Delta F))} - \sum_{i=1}^{N_R} \frac{1}{1 + \exp(-\beta(M + W_i - \Delta F))} \quad (27)$$

is equal to zero, where $M = kT \ln(N_F/N_R)$ and N_F and N_R are the number of realizations of the forward and reverse processes, respectively.⁴⁵

We compared the bias ($B(N)$) and the variance ($\text{Var}(N)$) of the free-energy estimates obtained by using the forward transition as described in the previous section with the estimates obtained by combining the forward and reverse processes. For this, after computing the forward and reverse processes we calculated the ΔF values for every step using the BAR (eq 27) and then added the ΔF values. As an extra control step, given the $N \times M$ matrix that describes the work done per step for the forward process, we built the work probability distribution by randomly selecting individual ordered steps from the N_F trajectories until a new trajectory was built. By using this approach, an ensemble of 10^5 trajectories was generated and used to construct the work probability distribution $P_F(W)$. The same procedure was used to build the work probability distribution for the reverse process $P_R(-W)$. It is important to verify that both distributions intercept. Also, when the CFT is used (eq 26), ΔF is determined as the intersection of the two distributions. The described BAR and CFT approaches give equivalent free-energy estimates, but the BAR approach is computationally more efficient but the CFT method allows us to visualize the

distributions graphically and should be used as a control of the results obtained.

The bias ($B(N)$) and the variance ($\text{Var}(N)$) of the free-energy estimates were computed for the transition between two offset harmonic wells of the same curvature.⁴² For this system, both $P_F(W)$ and $P_R(-W)$ are Gaussian distributions. Three different variances ($\sigma_T^2 = (kT)^2$, $8(kT)^2$, and $16(kT)^2$) were considered, assuming the same variance for the forward and reverse processes.

Bias and Variance for Non-Gaussian Work Fluctuations. To investigate the accuracy of the MSTC method for systems with non-Gaussian work fluctuations, we used the physically motivated model for the adiabatic compression of a dilute gas of interacting classical molecules.³⁶ This model is an example of a thermodynamic process for which non-Gaussian work fluctuations can be computed exactly from microscopic principles. For this system, the work probability distribution is

$$P(W) = \frac{\beta}{|\alpha_c| \Gamma(k)} \left(\frac{\beta|W|}{\alpha_c} \right)^{k-1} e^{\beta|W|/\alpha_c} \theta(\alpha_c W) \quad (28)$$

where $\alpha_c = (V_0/V_1)^{2/d} - 1$, with V_0 and V_1 being the initial and final volumes, respectively, d being the dimensionality of the space (for example, if we are describing the behavior of a particle in 3D space, $d = 3$), $k = Nd/2$ (N is the number of particles), and $\Gamma(k)$ is the gamma function. The step unit function $\theta(\cdot)$ guarantees that W is positive for compression ($\alpha_c > 0$) and negative for expansion ($\alpha_c < 0$). Equation 28 corresponds to a gamma distribution with shape parameter $k = Nd/2$ and scale parameter $s = |\alpha_c|/\beta$. Different parameter combinations were considered to produce forward distributions with the difference variance ($\sigma_T^2 = (kT)^2$, $8(kT)^2$, and $16(kT)^2$).

RESULTS

We compared the bias of $\Delta \hat{F}$ estimations for processes with different total variances ($\sigma_T^2 = (kT)^2$, $2(kT)^2$, $4(kT)^2$, $8(kT)^2$, $16(kT)^2$ and $32(kT)^2$) by the MSTC method and by Jarzynski averaging over the one-step process (OSP, Figure 2) when $P(W) \sim \mathcal{N}(\mu, \sigma_T^2)$. In all cases, the MSTC method has a lower bias than the OSP averaging. In Figure 2, the continuous lines are the values of eqs 22 and 23 for OSP and MSTC estimates, respectively. Good agreement is observed between values obtained from the analytically derived expressions and numerical estimations (circles and triangles) for both the small- N and large- N limits. For processes with low variance (for example, $\sigma_T^2 = (kT)^2$), the differences in the bias are small in both the small- and large- N regimes. For processes with higher variances, OSP averaging always shows a higher bias in both regimes when compared to the MSTC method. For example, if a threshold bias (i.e., maximum allowed error) of $0.3kT$ is considered to be acceptable when $\sigma_T^2 = 8(kT)^2$, then the number of trajectories needed using OSP averaging is 300 whereas it is 20 using the MSTC method, a 15-fold improvement. For the same bias threshold, when $\sigma_T^2 = 16(kT)^2$ the number of trajectories needed using the MSTC method is 120 whereas it is 140 000 for OSP averaging, a 1200-fold improvement. Similar behavior is observed if the work probability distribution ($P(W)$) is assumed to be described by a gamma distribution instead of a normal distribution (see below).

The variance of the free-energy estimates ($\text{Var}(N)$) is also always lower when the MSTC averaging method is used, compared to OSP averaging (Figure 3). The behavior is similar

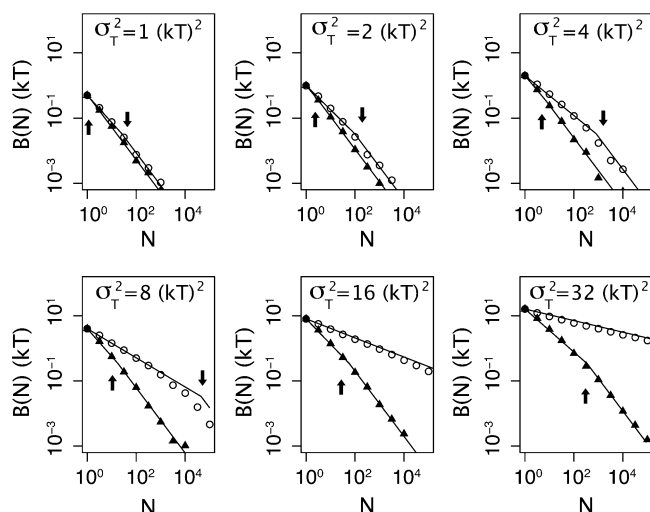


Figure 2. Free-energy bias for normal work probability distributions of different variances (log–log plot). Open circles represent numerically simulated values using JE averaging for OSP, and black triangles show values obtained using the MSTC methods. Lines represent analytically derived expressions for the bias using constants $C_{\text{OSP}} = 10$ and $C_{\text{MSTC}} = 10$. Arrows show where the small- and large- N regimes intersect. When arrows are not present, the intersection is not in the plot range.

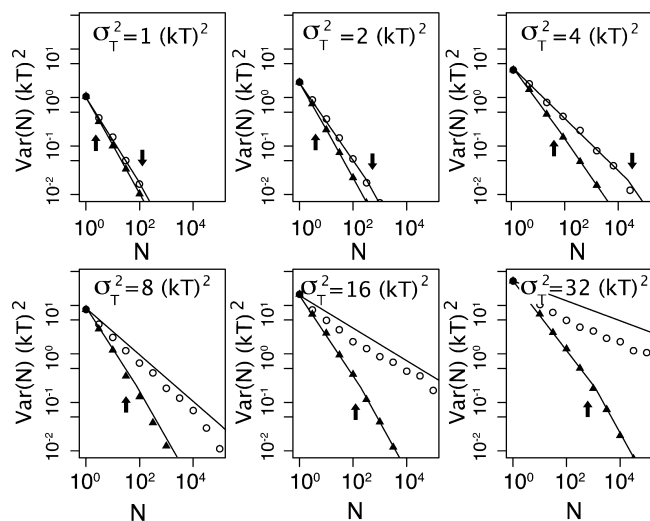


Figure 3. Variance of the free-energy estimates ($\text{Var}(N)$) for normal work probability distributions of different variances (σ_T^2) (log–log plot). Open circles and black triangles represent numerically simulated values using JE averaging for OSP and the MSTC methods, respectively. Lines represent analytically derived expressions for the variance using constant $C_v = 50$ (eq 25). Arrows show where the small- and large- N regimes intersect. When arrows are not present, the intersection is not in the range of the plot.

to that described for the bias. If a threshold variance of $0.3(kT)^2$ is used as the acceptance criterion, when $\sigma_T^2 = 8(kT)^2$ the number of trajectories needed when using OSP averaging is 700 whereas only 40 are needed when using the MSTC method, an 18-fold improvement. When $\sigma_T^2 = 16(kT)^2$, the number of trajectories needed using the MSTC method is 70, whereas 70000 are needed for JE averaging, a 1000-fold improvement. Figure 3 also shows that when the underlying variance of the process is high (i.e., $\sigma_T^2 = 16(kT)^2$ or higher) the analytically derived variances overestimate the variance on the ΔF estimates, which is consistent with the results of Gore et al.³⁹

Up to this point, we have assumed that all of the steps in a trajectory are described by the same work distribution (i.e., they all have the same variance (σ_s^2)). Using numerical simulations, we investigated the effect on the bias and the variance of ΔF when the $P(W)$ in one step has a higher variance than the rest but the total variance associated with the process is kept constant. The calculations show that when using MSTC averaging the improvement in reducing the bias ($B(N)$) and the variance ($\text{Var}(N)$) is maximal when all of the steps have the same variance (σ_s^2 , Figure 4). Therefore, to maximize the

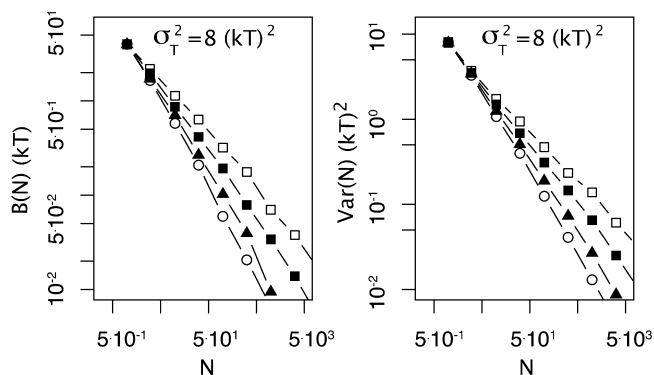


Figure 4. Free-energy bias and variance for normal work probability distributions using MSTC averaging (log–log plot). Calculations were made for trajectories of $M = 10$ steps, $\sigma_T^2 = 8(kT)^2$, and with one step having a different variance. (O) All steps have the same variance (10%); (▲) one step has 40% of the variance; (■) one step has 60% of the variance; and (□) one step has 80% of the variance.

accuracy in the free-energy estimates, the MSTC method should be implemented in such a way that all steps have a similar variance. For example, when implementing MSTC via MD simulations, if the reaction coordinate is divided into equal steps then the portions of the trajectory that cross sections of the energy landscape that are more rugged show a higher variance with respect to the work values (i.e., sampling a transition step). Therefore, a way to improve the accuracy of the MSTC method is to divide the reaction coordinate into steps of different lengths to enhance the sampling in regions with a high work variance. A simple protocol would involve starting with equal-sized steps. After a few trajectories are computed, if one or more steps show an unusually large variance, then the computation for those steps will be repeated by adding intermediate steps (i.e., dividing each step into two steps).

We also compared the accuracy of the free-energy estimates obtained from the forward work distributions (Figures 2 and 3) with the results obtained when data from the forward and reverse processes is available. As described above, we used two different models that allowed us to explore the behavior of the free-energy estimates for cases in which the work probability distribution was Gaussian or gamma. In both cases studied (Figures 5 and 6), free-energy estimates are more accurate when information about the forward and reverse processes is taken into consideration. This is shown as a reduction in the bias and variance (Figures 5 and 6). Of course, when the forward and reverse transitions are available the effective number of trajectories computed is $2N$ instead of N ; therefore, the advantage of computing the forward and reverse transitions is greater when the variance associated with the overall process is larger.

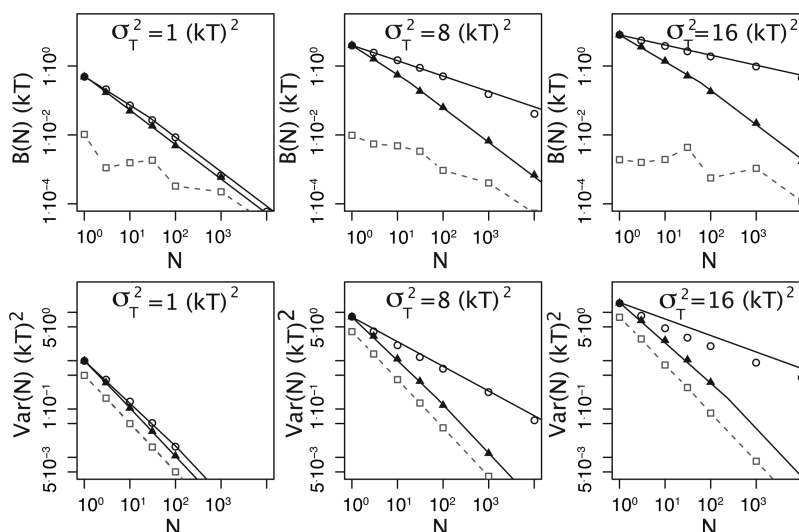


Figure 5. Free-energy bias and variance for normal work probability distributions when OSP or MSTC averaging is applied over the forward trajectories ($-\circ-$, OSP averaging; \blacktriangle , MSTC averaging) and when the forward and reverse trajectories are combined using the BAR method ($-\square-$). Forward trajectories computed for $\sigma_T^2 = (kT)^2$, $8(kT)^2$, and $16(kT)^2$ are shown.

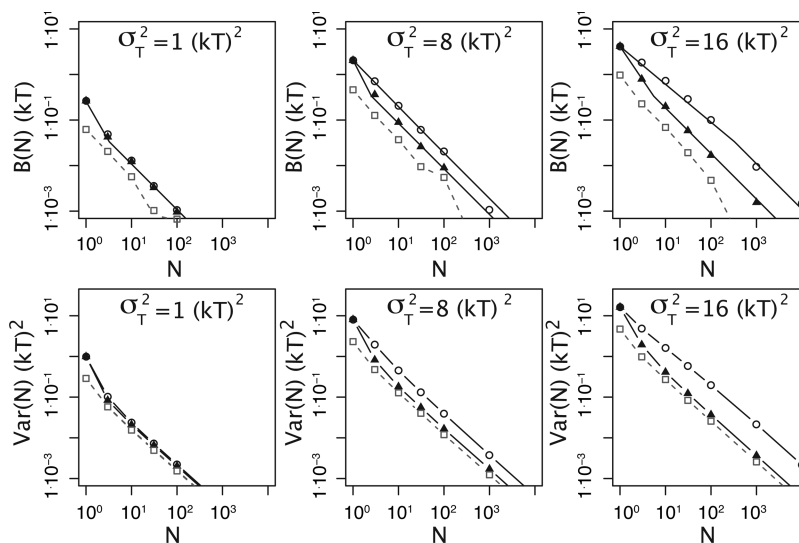


Figure 6. Free-energy bias and variance for gamma work probability distributions when OSP or MSTC averaging are applied over the forward trajectories ($-\circ-$, OSP averaging; \blacktriangle , MSTC averaging) and when forward and reverse trajectories are combined using the BAR method ($-\square-$). Forward trajectories computed for $\sigma_T^2 = (kT)^2$, $8(kT)^2$, and $16(kT)^2$ are shown.

As defined in this article, the MSTC and OSP processes are compared by assuming the same total time; therefore, they can be compared in terms of the computational efficiency. In addition, we assumed that in the procedures used to obtain the work values there is no way to decide when the dissipation ends and the equilibration starts: the system is under force all the time. Nevertheless, the tools presented in this article are general and can be used to compare other protocols of the MSTC and OSP averaging methods. The comparisons were set up by assuming that the total variance and the distribution obtained with both methods are comparable but can also be used in situations where the MSTC and OSP have different total variances. Depending on the system that is being studied, it is possible that the total variance for the two procedures is different. In cases where the variance in the work values obtained using the OSP process is lower, performing OSP simulations could yield more accurate results depending on the variance difference between the two methods.

In general, given the assumption that the distribution of work values for an OSP and MSTC protocol are comparable, the MSTC method will give more accurate results. The stepwise way that the MSTC method has been set favors this method over the OSP averaging at three different levels: (1) Equilibrating the system at intermediate stages prevents the system from going very far from equilibrium and accessing very low probability states. Sampling these very low probability states will produce a work probability distribution that is not representative of the transition, compromising the accuracy of the results. Also, if the system is driven far from equilibrium and then as it is further driven along the reaction coordinate, the system will constantly be catching up and any mechanistic conclusions obtained from these simulations will not be representative of the process being studied. (2) If the process were to be performed in one step (at the same velocity as that used for the MSTC method) without intermediate equilibration, then the variance of the work values would be higher. As

has been shown, what determines the number of trajectories needed to compute an accurate free-energy difference is the variance of the work values. In this context, using the OSP method will result in needing more trajectories to achieve the same accuracy. (3) Combining trajectories results in a net effect of having more trajectories, which effectively reduces the bias and variance of the free-energy estimates. Therefore, even if the intermediate equilibrations after each step increase the computational time, the gain in computational efficiency will still be determined by the variance of the work values. In addition, dividing the trajectories into steps has the extra advantage of making the computation of the reverse transition by reversing each of the steps in the forward process feasible.

In MD simulations, the fact that the MSTC method divides the reaction coordinate into steps can be exploited to obtain forward and reverse work probability distributions. For a process connecting states A and B, one direction would most likely be much easier to compute by MD simulations in terms of the time required to achieve equilibrium. One classic example is protein folding: although unfolding a protein by pulling apart the ends is easily achievable, it is extremely unlikely that the protein will refold by reversing the pulling path. Nevertheless, when a small perturbation such as an MSTC step is applied in the forward direction, it will be possible for the system to return to its initial configuration if the perturbation is reversed. Therefore, even if the difference in accuracy between the MSTC averaging and OSP averaging when the forward and reverse transitions are combined might not seem very significant (Figures 5 and 6), in most practical situations the reverse transition cannot be computed as a single-step process. When MD simulations are performed using the MSTC protocol, because equilibrium conformations are saved at the end of each step in the forward realization, reverse simulations can be started at these structures to reverse the process. We have successfully used this method for two systems: helix unfolding and ligand binding.^{43,44}

Relation of the MSTC Method to the Potential of Mean Force (PMF). The free-energy profile or potential of mean force (PMF, Φ) as a function of a reaction coordinate ξ is given by

$$\exp[-\beta\Phi(\xi')] = \int d\mathbf{r} d\mathbf{p} \delta(\xi(\mathbf{r}) - \xi') \exp[-\beta H(\mathbf{r}, \mathbf{p})] \quad (29)$$

where $\beta = 1/kT$, δ is a Dirac delta function, and H is the Hamiltonian. To compute this quantity using MD simulations, it is necessary to explore the system along the reaction coordinate. This usually relies on umbrella sampling (US)¹⁶ to increase the sampling of high-energy states by introducing a biasing potential (V_B , usually a harmonic potential) at different positions (ξ_i) along the reaction coordinate (ξ) and carrying out MD simulations using a modified Hamiltonian

$$H_i(\mathbf{q}) = H(\mathbf{q}) + V_B(\xi - \xi_i) \quad (30)$$

Another method of estimating the PMF is steered molecular dynamics (SMD),^{21,46} where a guiding potential V_λ is used instead of a biasing potential. Here, parameter λ is typically changed at constant velocity, and the modified Hamiltonian is a function of λ :

$$H_\lambda(\mathbf{q}) = H(\mathbf{q}) + V_\lambda(\lambda) \quad (31)$$

Several heuristic procedures have been implemented to remove the effects of the biasing potential on the distribution of

states accessed in this type of MD simulation. The removal of the bias requires that the MD runs be carried out with the biasing potential centered at closely spaced points ξ_i along reaction coordinate ξ . In the weighted histogram analysis method (WHAM),⁴⁷ the spacing between points should be such that the histograms (built by counting the number of accessed states at each value of the coordinate ξ) from adjacent windows (centered at the different ξ_i positions of the biasing potential) have enough overlap to allow the values of $H(\mathbf{q})$ in the overlapped regions to be matched. In addition, different methods have been developed to obtain the PMF using Jarzynski's equality for transitions in one direction^{20,21,24,46} and for bidirectional transitions.^{48–51}

MD simulations computed at closely spaced centers of the biasing potential can be used to estimate the free-energy differences between adjacent steps using Jarzynski's equation (eq 2) if the work necessary to take the system from one biasing point to the next is recorded. According to this equation, the change in free energy between states i and $i + 1$ can be estimated from the relation

$$\Delta F_{i,i+1} = -\beta^{-1} \log \langle e^{-\beta W_{ij}} \rangle_j \quad (32)$$

where $\langle e^{-\beta W_{ij}} \rangle_j$ is an average over repetitions of the process that take the system from state i to state $i + 1$ by doing work W_{ij} ($j = 1, \dots, N$, where N is the number of repetitions). This relation is valid even if the process is carried out irreversibly as long as in all of the repetitions the initial and final states i and $i + 1$ are equilibrium states.

Two modifications of the MD simulations used in PMF calculations are needed to use Jarzynski's equation. First, before the biasing potential centered at ξ_{i+1} is introduced, the system must be equilibrated with the biasing potential centered at ξ_i . Then, a new spring with its equilibrium position at ξ_{i+1} is set instantaneously. This would allow us to compute the work that takes the system from i to $i + 1$ over the evolution of the system as it is being pulled between the two states by the biasing potential centered at $i + 1$.⁵² Second, this computation has to be repeated for each i and $i + 1$ pair enough times, after equilibrating at each state, to provide an accurate value of the exponential average.

MSTC Method Applied to MD Simulations. As already mentioned, accurate numerical estimations of free-energy differences require a well-sampled work probability distribution ($P(W)$). The MSTC method has previously been used in MD studies of the amino acid helix propensity⁴³ and disaccharide binding to galectin-1.⁴⁴ In those cases, results obtained by the MSTC method reproduce the experimental data accurately. This suggests that the MSTC method can be used to extract as much information as possible from a limited set of trajectories obtained by MD simulations.

To apply the MSTC method using MD simulations, a case-by-case system-dependent protocol has to be defined. The protocol requires the definition of the system reaction coordinate and the velocity at which the system is driven between predefined states A and B. To move the system along the reaction coordinate, it is necessary to apply a force using a soft constraint. For example, when calculating free-energy differences for ligand binding,⁴⁴ a dummy atom located in the center of mass was tethered to the ligand. This dummy atom was displaced a small distance ($\ll 1$ Å), and the system was relaxed for a short time (there is no need for the system to reach equilibrium at this stage) with the dummy atom fixed.

Hereafter, these are referred to as small steps. After K small steps have been completed, the system needs to be equilibrated. These are referred to as long steps. The ratio between the distance displaced in every small step and the equilibration time will be referred to as the pulling velocity. This quantity will influence the fluctuations in the work values: as the pulling velocity is increased, the variance in the work probability distribution will also increase. Another quantity that will affect the fluctuations in the work values is the tether force constant: increasing the force constant decreases the fluctuation in the work values. The extent to which the pulling velocity and the force constant affect the fluctuations is system-dependent, and a proper determination of both should be done by exploring different combinations. In addition, it is important to equilibrate between long steps. To determine the equilibration times, it is necessary to guarantee the decorrelation of the velocities, energy, and temperature in addition to allowing the system to equilibrate to the new equilibrium position.

The exponential averaging used in Jarzynski's equation heavily weights rare events. Hence, a high variance in the work values increases the probability of accessing these rare events. However, it has been shown that a higher variance will increase the systematic bias and variance of the free-energy estimates. A proper MSTC protocol balances the fluctuations in such a way that the number of trajectories needed is accessible by MD simulations. Results presented in this work provide important clues for estimating the number of trajectories needed for a given calculation.

SUMMARY AND CONCLUSIONS

We previously introduced the MSTC method to calculate free-energy differences using MD simulations.^{43,44} The method is based on the use of Jarzynski's equation^{26,27} to evaluate the free-energy difference between two states by dividing the transition into steps that are later combined to generate a path ensemble. If the system is equilibrated with identical external conditions, then microstates at the end of each step belong to the same ensemble, making trajectory combination possible. Therefore, a combinatorial process can be used to generate a representative work distribution from a small number of computed trajectories. This combination of processes between equivalent microstates gives rise to a work distribution for the transition, which is used to estimate the free energy. In addition, we have proven that this is equivalent to computing ΔF per step and then adding them.

The MSTC method can be used to obtain accurate free-energy estimates with significantly lower bias and variance when compared to simple exponential averaging over one-step trajectories given that the variances in the work values are comparable. In addition, the fact that the transition is divided into steps makes it possible to compute the reverse process. The availability of the forward and reverse processes can be used to estimate the free-energy difference when used in combination with the BAR⁴⁵ method. When the BAR method is used to estimate ΔF , the values obtained are more accurate as judged by comparing the bias and variance of the same estimates to those obtained using only the forward process. Therefore, when MD simulations are performed, the MSTC method makes it possible to reverse the transition and to obtain more accurate estimates of ΔF by combining the forward and reverse work probability distributions.

Another major advantage of the MSTC method is that it is coupled to a physical path. Hence, information gathered along

this path can be used to obtain mechanistic information about the process under study. Nevertheless, this is also a limitation of the method because choosing an appropriate reaction coordinate can be challenging. Also, as the system is driven out of equilibrium it is possible to explore the atomic detail of intermediate metastable states present between the initial and final states. This feature is important for extracting mechanistic information about how the interactions that stabilize the initial states change into those that stabilize the final state. As described, the MSTC method is based on a sequence of displacements followed by equilibrations in which the system is allowed to relax. For every trajectory, these equilibrium intermediates can be analyzed to obtain physical properties of the system and mechanistic information about the process. This differs from other methods such as steered molecular dynamics (SMD) where the system is pulled continuously, making mechanistic conclusions difficult to justify because the system is constantly catching up with the applied perturbation. This could be overcome if an ensemble of SMD trajectories is considered as proposed by Crooks.²⁸

The MSTC method is general enough to be applied to different processes such as the estimation of differences in the binding affinity, the unfolding of proteins with a small number of mutations, the unfolding of RNAs with individual substitutions, and other similar processes. The computation of free energies using the MSTC method can provide new insight into two different and complementary situations: (1) They can be used to obtain estimates of the free energies when experimental data are not available. (2) When experimental data are available, the computed values of the free energy can be compared to the experimental values to validate the computations. If the computed values prove to be accurate, then an analysis of the trajectories provides mechanistic insight into the details of binding and unbinding.

ASSOCIATED CONTENT

Supporting Information

Bias and variance for gamma work probability distributions. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: mamzel@jhmi.edu.

Present Address

[†]Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742, United States.

Notes

The authors declare no competing financial interest.

REFERENCES

- (1) Shea, J. E.; Brooks, C. L. *Annu. Rev. Phys. Chem.* **2001**, *52*, 499.
- (2) Simonson, T.; Archontis, G.; Karplus, M. *Acc. Chem. Res.* **2002**, *35*, 430.
- (3) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562.
- (4) Bai, D.; Elber, R. J. *Chem. Theory Comput.* **2006**, *2*, 484.
- (5) Wyczalkowski, M. A.; Pappu, R. V. *Phys. Rev. E* **2008**, *77*, 026104.
- (6) Park, S.; Lau, A. Y.; Roux, B. *J. Chem. Phys.* **2008**, *129*, 134102.
- (7) Deng, Y. Q.; Roux, B. *J. Phys. Chem. B* **2009**, *113*, 2234.
- (8) Christ, C. D.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2010**, *31*, 1569.

- (9) Zhuravlev, P. I.; Wu, S.; Potoyan, D. A.; Rubinstein, M.; Papoian, G. A. *Methods* **2010**, *52*, 115.
- (10) Wereszczynski, J.; McCammon, J. A. *Q. Rev. Biophys.* **2011**, *1*.
- (11) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150.
- (12) Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1989**, *90*, 2460.
- (13) Wood, R. H.; Muhlbauer, W. C. F.; Thompson, P. T. J. *Phys. Chem.* **1991**, *95*, 6670.
- (14) Aqvist, J. *J. Phys. Chem.* **1990**, *94*, 8021.
- (15) Chipot, C.; Kollman, P. A.; Pearlman, D. A. *J. Comput. Chem.* **1996**, *17*, 1112.
- (16) Torrie, G. M.; Valleau, J. P. *Chem. Phys. Lett.* **1974**, *28*, 578.
- (17) Straatsma, T. P.; Berendsen, H. J. C. *J. Chem. Phys.* **1988**, *89*, 5876.
- (18) Straatsma, T. P.; McCammon, J. A. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407.
- (19) Hummer, G. *J. Chem. Phys.* **2001**, *114*, 7330.
- (20) Hummer, G.; Szabo, A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3658.
- (21) Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. *J. Chem. Phys.* **2003**, *119*, 3559.
- (22) Atilgan, E.; Sun, S. X. *J. Chem. Phys.* **2004**, *121*, 10392.
- (23) Bastug, T.; Chen, P. C.; Patra, S. M.; Kuyucak, S. *J. Chem. Phys.* **2008**, *128*, 155104.
- (24) Hummer, G.; Szabo, A. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 21441.
- (25) Chipot, Ch.; Pohorille, A., Eds. *Free Energy Calculations: Theory and Applications in Chemistry and Biology*; Springer Series in Chemical Physics; Springer: Berlin, 2007.
- (26) Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690.
- (27) Jarzynski, C. *Phys. Rev. E* **1997**, *56*, 5018.
- (28) Crooks, G. E. *Phys. Rev. E* **2000**, *61*, 2361.
- (29) Evans, D. J. *Mol. Phys.* **2003**, *101*, 1551.
- (30) Chernyak, V.; Chertkov, M.; Jarzynski, C. *Phys. Rev. E* **2005**, *71*, 025102.
- (31) Liphardt, J.; Onoa, B.; Smith, S. B.; Tinoco, I. J.; Bustamante, C. *Science* **2001**, *292*, 733.
- (32) Wang, G. M.; Sevick, E. M.; Mittag, E.; Searles, D. J.; Evans, D. J. *Phys. Rev. Lett.* **2002**, *89*, 050601.
- (33) Trepagnier, E. H.; Jarzynski, C.; Ritort, F.; Crooks, G. E.; Bustamante, C. J.; Liphardt, J. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15038.
- (34) Collin, D.; Ritort, F.; Jarzynski, C.; Smith, S. B.; Tinoco, I., J.; Bustamante, C. *Nature* **2005**, *437*, 231.
- (35) Imparato, A.; Peliti, L. *Phys. Rev. E* **2005**, *72*, 046114.
- (36) Crooks, G. E.; Jarzynski, C. *Phys. Rev. E* **2007**, *75*, 021116.
- (37) Calderon, C. P.; Chelli, R. *J. Chem. Phys.* **2008**, *128*, 145103.
- (38) Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235.
- (39) Gore, J.; Ritort, F.; Bustamante, C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12564.
- (40) Zuckerman, D. M.; Woolf, T. B. *Phys. Rev. Lett.* **2002**, *89*, 180602.
- (41) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. *Phys. Rev. Lett.* **2003**, *91*, 140601.
- (42) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 144107.
- (43) Echeverria, I.; Amzel, L. M. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1302.
- (44) Echeverria, I.; Amzel, L. M. *Biophys. J.* **2011**, *100*, 2283.
- (45) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245.
- (46) Park, S.; Schulten, K. *J. Chem. Phys.* **2004**, *120*, 5946.
- (47) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1658.
- (48) Chelli, R.; Marsili, S.; Procacci, P. *Phys. Rev. E* **2008**, *77*, 031104.
- (49) Chelli, R.; Procacci, P. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1152.
- (50) Nicolini, P.; Procacci, P.; Chelli, R. *J. Phys. Chem. B* **2010**, *114*, 9546.
- (51) Minh, D. D. L.; Adib, A. B. *Phys. Rev. Lett.* **2008**, *100*, 180602.
- (52) Adib, A. B. *J. Chem. Phys.* **2009**, *130*, 247101.