

Calculating potentials of mean force from steered molecular dynamics simulations

Sanghyun Park and Klaus Schulten

Beckman Institute and Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

(Received 3 December 2003; accepted 7 January 2004)

Steered molecular dynamics (SMD) permits efficient investigations of molecular processes by focusing on selected degrees of freedom. We explain how one can, in the framework of SMD, employ Jarzynski's equality (also known as the nonequilibrium work relation) to calculate potentials of mean force (PMF). We outline the theory that serves this purpose and connects nonequilibrium processes (such as SMD simulations) with equilibrium properties (such as the PMF). We review the derivation of Jarzynski's equality, generalize it to isobaric–isothermal processes, and discuss its implications in relation to the second law of thermodynamics and computer simulations. In the relevant regime of steering by means of stiff springs, we demonstrate that the work on the system is Gaussian-distributed regardless of the speed of the process simulated. In this case, the cumulant expansion of Jarzynski's equality can be safely terminated at second order. We illustrate the PMF calculation method for an exemplary simulation and demonstrate the Gaussian nature of the resulting work distribution. © 2004 American Institute of Physics. [DOI: 10.1063/1.1651473]

I. INTRODUCTION

A key goal of the study of biomolecular systems is to identify the physical mechanisms establishing their functions. In a typical investigation of a respective molecular process, the *reaction path*, along which the process proceeds in the configuration space, is identified or hypothesized and the progress of the process is described by the *reaction coordinate*.¹ The *potential of mean force* (PMF) plays an important role in such investigations. PMF is basically the free energy profile along the reaction coordinate and is determined through the Boltzmann-weighted average over all degrees of freedom other than the reaction coordinate. PMF not only succinctly captures the energetics of the process studied, but also provides an essential ingredient for further modeling of the process; with all the other degrees of freedom averaged out, the motion along the reaction coordinate is well approximated as a diffusive motion on the effective potential identified as the PMF.

Molecular dynamics is a simulation method widely applied to biomolecular systems.² However, today's molecular dynamics simulations are limited to the nanosecond time scale which is seldom long enough to observe relevant processes. *Steered molecular dynamics* (SMD) therefore applies external steering forces in the right direction to accelerate processes that otherwise, due to energy barriers, are too slow. SMD, reviewed in Refs. 3 and 4, has been widely used to investigate mechanical functions of proteins such as stretching of extracellular matrix or muscle proteins^{5–7} and binding/unbinding of protein–substrate complexes or adhesion proteins.^{8,9} A typical SMD simulation steers a system by applying a constraint (e.g., a harmonic potential) that moves along a prescribed path in the configuration space.

As SMD is an effective method to explore molecular processes, it is desirable to calculate within its framework

PMFs. However, a SMD simulation is a *nonequilibrium* process, whereas PMF is an *equilibrium* property. Therefore a theory is needed that connects equilibrium and nonequilibrium. Such a theory has become available through recent advances in nonequilibrium statistical mechanics, especially through the discovery of Jarzynski's equality.¹⁰ These advances permit one to extract equilibrium properties from nonequilibrium processes, but in practice efficient and convenient methods are required.

Jarzynski's equality is an exact relation between free energy differences and the work done through nonequilibrium processes. Since its first report in 1997,^{10,11} Jarzynski's equality has been a subject of intensive study. The relation with the fluctuation theorems was elucidated by Crooks¹² and by Jarzynski.¹³ Hatano and Sasa¹⁴ generalized Jarzynski's equality to transformations between steady states based on the steady state thermodynamics of Oono and Paniconi.¹⁵ And recently Liphardt *et al.*¹⁶ tested Jarzynski's equality in an experiment of RNA stretching.

Jarzynski's equality finds a natural application in the calculation of free energy or PMF from computer simulations or experiments.^{17–22} Particularly, it provides the basis for the method presented in this article for calculating PMFs from SMD simulations. The method has been applied to the investigations of protein functions such as glycerol conduction through the membrane channel GlpF (Ref. 23) and ammonia conduction through HisF.²⁴ In a benchmark study using the helix–coil transition of deca-alanine as an exemplary system, the accuracy of the approximations based on the cumulant expansion was examined and compared to the traditional method of umbrella sampling.²⁵

This article is concerned with theoretical and practical issues regarding the method of PMF calculation from SMD simulations. Section II reviews and discusses the theoretical

background. In Sec. III the method is presented and related practical issues are discussed, in particular the efficiency and convenience of the method. In Sec. IV the method is illustrated with an exemplary SMD simulation and the Gaussian nature of the work distribution is demonstrated.

II. THEORY OF SYSTEMS DRIVEN AWAY FROM EQUILIBRIUM

Thermodynamics is concerned with states of matter and transformations between them. Statistical mechanics started with the aim of explaining the laws of thermodynamics based on the atomic picture of matter. It has been successful with systems at equilibrium, but not quite so with nonequilibrium ones. Most of the development in statistical mechanics for nonequilibrium states has been limited to near-equilibrium (linear response) regimes. However, recently there has been some further progress through the proof of theorems concerning far-from-equilibrium states: the transient fluctuation theorem,²⁶ the steady-state fluctuation theorem,²⁷ and Jarzynski's equality also known as the nonequilibrium work relation.¹⁰

This section deals with Jarzynski's equality, which is the basis of the PMF calculation featured in this article. Two different derivations, one for Hamiltonian systems and the other for stochastic systems, are presented. Jarzynski's equality is generalized to isobaric–isothermal processes, and the relationship with the second law of thermodynamics is discussed. Finally, it is demonstrated that Jarzynski's equality can be applied to computer simulations, in particular the isobaric–isothermal molecular dynamics simulation using the Langevin piston method.²⁸

A. Jarzynski's equality

Jarzynski's equality is concerned with thermostated (in contact with heat baths) systems that begin in equilibrium and subsequently are driven away from equilibrium. Let us consider a system in contact with a heat bath at temperature T . Suppose the equilibrium states of the system are specified by (T, λ) , where λ is a parameter that can be controlled externally. Initially λ is, say, zero and the system is in the equilibrium state $(T, 0)$. The parameter λ is then changed, say, up to Λ . Over the entire course of this process, the system is kept in contact with the heat bath. Let W be the external work done on the system during the process of increasing λ . We imagine to repeat the process many times. Jarzynski¹⁰ discovered that the Helmholtz free energy difference ΔF between the equilibrium states, (T, Λ) and $(T, 0)$, is related to the work W as

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

where $\beta = 1/k_B T$ is the inverse temperature and k_B is the Boltzmann constant. The average $\langle \cdot \rangle$ is over repeated realizations of the process.

Although in general final states of the system will not be in equilibrium, one can fix λ at Λ and wait for the system to relax to the equilibrium state (T, Λ) . During the relaxation, no external work is done. Therefore, Jarzynski's equality can be stated in terms of transformations between equilibrium states: *for a transformation between two equilibrium states*

of a system in contact with a heat bath at temperature T , the work W done on the system during the transformation and the Helmholtz free energy difference ΔF between the two equilibrium states satisfy Eq. (1).

As shown in Secs. II B and II C, Jarzynski's equality applies to a broad range of processes. The system under consideration may be microscopic or macroscopic, and the parameter λ may be, but is not limited to, a thermodynamic variable (intensive or extensive). In particular, Jarzynski's equality does apply to traditional thermodynamic processes such as the isothermal expansion of an ideal gas. In the latter case the parameter λ is the volume of the gas. Most remarkably, Jarzynski's equality holds regardless of the speed of the process.

B. Jarzynski's equality for Hamiltonian systems

Jarzynski's equality was first derived for Hamiltonian systems,¹⁰ as outlined in the following. Consider a classical mechanical system in contact with a heat bath of constant temperature T . Let us label the system \mathcal{S} , the bath \mathcal{B} , and the compound of the two \mathcal{SB} . The compound \mathcal{SB} is thermally isolated and evolves according to Hamiltonian dynamics. Assuming that the surface energy (or the interaction energy) between \mathcal{S} and \mathcal{B} is negligible, the Hamiltonian of \mathcal{SB} can be divided into the Hamiltonians of \mathcal{S} and \mathcal{B} :

$$H_\lambda^{\mathcal{SB}}(\Gamma, \Theta) = H_\lambda^{\mathcal{S}}(\Gamma) + H^{\mathcal{B}}(\Theta), \quad (2)$$

where Γ and Θ denote phases (positions and momenta) of \mathcal{S} and \mathcal{B} , respectively, and the Hamiltonian of \mathcal{S} depends on a parameter λ . The partition function, therefore, is factorized,

$$\begin{aligned} Z_\lambda^{\mathcal{SB}} &= \int d\Gamma d\Theta \exp[-\beta H_\lambda^{\mathcal{SB}}(\Gamma, \Theta)] \\ &= \int d\Gamma \exp[-\beta H_\lambda^{\mathcal{S}}(\Gamma)] \int d\Theta \exp[-\beta H^{\mathcal{B}}(\Theta)] \\ &= Z_\lambda^{\mathcal{S}} Z^{\mathcal{B}}. \end{aligned} \quad (3)$$

Now consider a process in which the system \mathcal{S} is initially in equilibrium with the bath \mathcal{B} and subsequently the parameter is changed from 0 at time 0 to Λ at time τ . The time evolution of \mathcal{SB} is determined by the time-dependent Hamiltonian $H_\lambda^{\mathcal{SB}}(\Gamma, \Theta)$, where the explicit time dependence solely comes from the dependence on λ . Let us denote the initial and final states by (Γ_0, Θ_0) and $(\Gamma_\tau, \Theta_\tau)$, respectively. Since \mathcal{SB} is thermally isolated, the distribution of its initial states would be best represented by a microcanonical ensemble. However, as \mathcal{SB} is a macroscopic system (even when \mathcal{S} is not), it is permissible to use a canonical ensemble instead. Accordingly, (Γ_0, Θ_0) may be sampled from the distribution

$$\frac{1}{Z_0^{\mathcal{SB}}} \exp[-\beta H_0^{\mathcal{SB}}(\Gamma_0, \Theta_0)]. \quad (4)$$

Because of energy conservation, the work done during the process must be equal to the increase in the energy of \mathcal{SB} ,

$$W = H_\Lambda^{\mathcal{SB}}(\Gamma_\tau, \Theta_\tau) - H_0^{\mathcal{SB}}(\Gamma_0, \Theta_0). \quad (5)$$

Thus the average of exponential work is written as

$$\begin{aligned} \langle e^{-\beta W} \rangle &= \int d\Gamma_0 d\Theta_0 \frac{1}{Z_0^{SB}} \exp[-\beta H_0^{SB}(\Gamma_0, \Theta_0)] \\ &\quad \times \exp\{-\beta[H_\Lambda^{SB}(\Gamma_\tau, \Theta_\tau) - H_0^{SB}(\Gamma_0, \Theta_0)]\} \\ &= \int d\Gamma_0 d\Theta_0 \frac{1}{Z_0^{SB}} \exp[-\beta H_\Lambda^{SB}(\Gamma_\tau, \Theta_\tau)]. \end{aligned} \quad (6)$$

The initial phase (Γ_0, Θ_0) and the final phase $(\Gamma_\tau, \Theta_\tau)$ are related by a one-to-one map: the final phase can be obtained from the initial phase by the forward time evolution of the Hamiltonian system, and the initial from the final by the backward time evolution. Therefore the integration variable can be transformed to $(\Gamma_\tau, \Theta_\tau)$, and according to Liouville's theorem the Jacobian of the transformation is unity,

$$\langle e^{-\beta W} \rangle = \int d\Gamma_\tau d\Theta_\tau \frac{1}{Z_0^{SB}} \exp[-\beta H_\Lambda^{SB}(\Gamma_\tau, \Theta_\tau)] = \frac{Z_\Lambda^{SB}}{Z_0^{SB}}. \quad (7)$$

Finally, using Eq. (3), we obtain Jarzynski's equality between the work W and the Helmholtz free energy difference ΔF^S of the system S ,

$$\langle e^{-\beta W} \rangle = \frac{Z_\Lambda^S}{Z_0^S} = \exp(-\beta \Delta F^S). \quad (8)$$

The use of the canonical ensemble [Eq. (4)] is a crucial point of the derivation. Jarzynski (Ref. 13, p. 100) argues that "the canonical ensemble should be viewed primarily as a computational convenience." This is justified as follows. The canonical average [Eq. (6)] can be expressed as a weighted integral over microcanonical averages,

$$\begin{aligned} \langle e^{-\beta W} \rangle_\beta^{\text{can}} &= \frac{1}{Z_0^{SB}} \int d\Gamma_0 d\Theta_0 \\ &\quad \times \exp[-\beta H_0^{SB}(\Gamma_0, \Theta_0)] e^{-\beta W} \\ &= \frac{1}{Z_0^{SB}} \int d\Gamma_0 d\Theta_0 \int dE \delta(H_0^{SB}(\Gamma_0, \Theta_0) - E) \\ &\quad \times \exp[-\beta H_0^{SB}(\Gamma_0, \Theta_0)] e^{-\beta W} \\ &= \frac{1}{Z_0^{SB}} \int dE \exp[-\beta E + S_0^{SB}(E)/k_B \\ &\quad + \log \langle e^{-\beta W} \rangle_E^{\text{mic}}], \end{aligned} \quad (9)$$

where $\langle e^{-\beta W} \rangle_\beta^{\text{can}}$ denotes the canonical average at temperature $T = 1/k_B \beta$, $\langle e^{-\beta W} \rangle_E^{\text{mic}}$ the microcanonical average at energy E , and $S_0^{SB}(E)$ the entropy of SB at energy E and $\lambda = 0$. The integral over E in Eq. (9) is dominated by the value of E that maximizes the integrand, namely that satisfies

$$-\beta + \frac{1}{k_B} \frac{\partial}{\partial E} S_0^{SB}(E) + \frac{\partial}{\partial E} \log \langle e^{-\beta W} \rangle_E^{\text{mic}} = 0. \quad (10)$$

Since the work W is done through the manipulation of the system S which is much smaller than the bath B , the work W must be much smaller than the energy scale of SB . Thus the third term in the left-hand side of Eq. (10) is negligible, and

the value of E that satisfies Eq. (10) is equal to the equilibrium energy of SB corresponding to the temperature T ; let us denote this energy by \tilde{E} . From Eq. (9), using $Z_0^{SB} = \exp[-\beta \tilde{E} + S_0^{SB}(\tilde{E})/k_B]$, we find

$$\langle e^{-\beta W} \rangle_\beta^{\text{can}} = \langle e^{-\beta W} \rangle_E^{\text{mic}}. \quad (11)$$

The use of the canonical ensemble instead of the microcanonical ensemble is therefore justified.

We have just derived Jarzynski's equality based on Hamiltonian dynamics. The derivation is surprisingly simple and depends only on fundamental properties of Hamiltonian dynamics, namely, energy conservation and Liouville's theorem. In the following we derive Jarzynski's equality in a different fashion.

C. Jarzynski's equality for stochastic systems

The dynamics of a system in contact with a heat bath is often described stochastically, without explicitly accounting for the degrees of freedom of the bath. Jarzynski's equality can be derived in this framework under two common assumptions, the Markov property and the balance condition. (In fact, in this framework Jarzynski's equality directly follows from the Feynman-Kac formula.) This type of derivation was first given in Ref. 11.

When the bath degrees of freedom are not explicitly taken into account, the dynamics of the system can be described only probabilistically, i.e., in terms of the probability distribution $f(\Gamma, t)$ for the microscopic state (or the phase) Γ of the system at time t . We assume that the time evolution of $f(\Gamma, t)$ is a *Markov process* described through

$$\partial_t f(\Gamma, t) = \mathcal{L}_{\lambda(t)} f(\Gamma, t). \quad (12)$$

The time evolution operator \mathcal{L}_λ depends on a parameter λ . We no longer need labels like S or B because we are now dealing with the system S only. We also assume that the equilibrium distribution

$$\Psi_\lambda(\Gamma) = \frac{1}{Z_\lambda} \exp[-\beta H_\lambda(\Gamma)] \quad (13)$$

is stationary under the time evolution

$$\mathcal{L}_\lambda \Psi_\lambda(\Gamma) = 0. \quad (14)$$

This is a weak form of detailed balance and, hence, will be referred to as the *balance condition*. The balance condition is a necessary condition for $f(\Gamma, t)$ to relax to $\Psi_\lambda(\Gamma)$ when λ is held constant. The system is initially in equilibrium corresponding to $\lambda = 0$, i.e.,

$$f(\Gamma, 0) = \Psi_0(\Gamma), \quad (15)$$

which provides an initial condition that accompanies Eq. (12).

For each realization of the process changing λ from 0 to Λ , a trajectory $\Gamma(t)$ is obtained. From a trajectory $\Gamma(t)$ we can calculate the work done on the system,

$$W[\Gamma(t)] = \int_0^\tau dt [\partial_t H_{\lambda(t)}(\Gamma)]_{\Gamma=\Gamma(t)}. \quad (16)$$

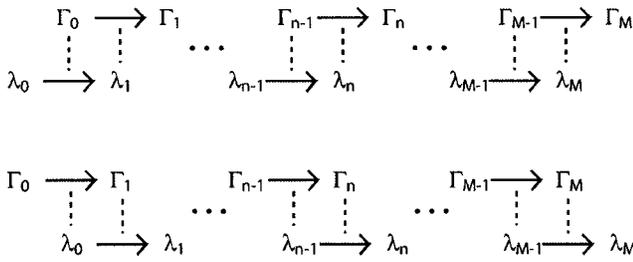


FIG. 1. Discretization schemes. Two schemes lead to the same final result. The scheme at the top is used here.

The work W is a functional that depends on the entire trajectory $\Gamma(t)$ for $0 < t < \tau$. This expression will be made clear when we discretize the process. The average $\langle e^{-\beta W} \rangle$ is then written as the path integral

$$\langle e^{-\beta W} \rangle = \int \mathcal{D}\Gamma(t) P[\Gamma(t)] e^{-\beta W[\Gamma(t)]}, \quad (17)$$

where the functional $P[\Gamma(t)]$ represents the probability for observing the trajectory $\Gamma(t)$. As the time evolution operator $\mathcal{L}_{\lambda(t)}$ and the initial probability $\Psi_0(\Gamma)$ completely determine the stochastic dynamics of the system, they must also determine $P[\Gamma(t)]$. This will also be made clear in the discretization.

The discretization scheme used is illustrated at the top of Fig. 1. Shown at the bottom of the figure is an alternative scheme. Both schemes lead to the same final result. Time is discretized as $t_n = n \delta t$ ($n = 0, 1, \dots, M$), with an infinitesimal interval $\delta t = \tau/M$. $\Gamma(t)$ and $\lambda(t)$ are discretized accordingly: $\Gamma_n = \Gamma(t_n)$ and $\lambda_n = \lambda(t_n)$, with $\lambda_0 = 0$ and $\lambda_M = \Lambda$. As shown in Fig. 1 the discretized process involves two alternating steps:

- (1) The parameter λ is externally changed from λ_{n-1} to λ_n while the system resides at Γ_{n-1} . During this step the amount of work, $H_{\lambda_n}(\Gamma_{n-1}) - H_{\lambda_{n-1}}(\Gamma_{n-1})$, is done on the system. The total amount of work done during the entire process is given as

$$W = \sum_{n=1}^M [H_{\lambda_n}(\Gamma_{n-1}) - H_{\lambda_{n-1}}(\Gamma_{n-1})], \quad (18)$$

which converges to Eq. (16) in the continuum limit ($M \rightarrow \infty$).

- (2) With λ fixed at λ_n , the system makes a transition from Γ_{n-1} to Γ_n due to its internal dynamics described through Eq. (12). No external work is done during this step. The change in the system energy, $H_{\lambda_n}(\Gamma_n) - H_{\lambda_n}(\Gamma_{n-1})$, can be attributed to the heat absorbed from the bath. During the entire process the system absorbs heat of the amount

$$Q = \sum_{n=1}^M [H_{\lambda_n}(\Gamma_n) - H_{\lambda_n}(\Gamma_{n-1})]. \quad (19)$$

We can easily check energy conservation,

$$W + Q = H_{\lambda_M}(\Gamma_M) - H_{\lambda_0}(\Gamma_0). \quad (20)$$

For each transition, $\Gamma_{n-1} \rightarrow \Gamma_n$, we denote by $R_{\lambda_n}(\Gamma_n | \Gamma_{n-1})$ the transition probability, i.e., the probability

that given its state Γ_{n-1} at t_{n-1} the system makes a transition to Γ_n at t_n . The probability to observe a certain discretized trajectory $(\Gamma_0, \dots, \Gamma_M)$ is then

$$P(\Gamma_0, \dots, \Gamma_M) = \Psi_{\lambda_0}(\Gamma_0) \prod_{n=1}^M R_{\lambda_n}(\Gamma_n | \Gamma_{n-1}). \quad (21)$$

Thus, in the discretized framework Eq. (17) becomes

$$\langle e^{-\beta W} \rangle = \int d\Gamma_0 \cdots d\Gamma_M \Psi_{\lambda_0}(\Gamma_0) \times \left[\prod_{n=1}^M R_{\lambda_n}(\Gamma_n | \Gamma_{n-1}) \right] e^{-\beta W}, \quad (22)$$

with W given as in Eq. (18). While the time evolution of the probability $f(\Gamma, t)$ is described in Eq. (12) in terms of the operator $\mathcal{L}_{\lambda(t)}$, in the discretized framework it is described in terms of the transition probability,

$$f(\Gamma_n, t_n) = \int d\Gamma_{n-1} R_{\lambda_n}(\Gamma_n | \Gamma_{n-1}) f(\Gamma_{n-1}, t_{n-1}). \quad (23)$$

The balance condition [Eq. (14)] means that the equilibrium distribution is stationary, which in the discretized framework implies

$$\Psi_{\lambda_n}(\Gamma_n) = \int d\Gamma_{n-1} R_{\lambda_n}(\Gamma_n | \Gamma_{n-1}) \Psi_{\lambda_n}(\Gamma_{n-1}). \quad (24)$$

Now we are ready to derive Jarzynski's equality. We start by writing $e^{-\beta W}$ in terms of the equilibrium distribution Ψ ,

$$e^{-\beta W} = \prod_{n=1}^M \frac{\exp[-\beta H_{\lambda_n}(\Gamma_{n-1})]}{\exp[-\beta H_{\lambda_{n-1}}(\Gamma_{n-1})]} = \prod_{n=1}^M \frac{Z_{\lambda_n} \Psi_{\lambda_n}(\Gamma_{n-1})}{Z_{\lambda_{n-1}} \Psi_{\lambda_{n-1}}(\Gamma_{n-1})} = \frac{Z_{\lambda_M}}{Z_{\lambda_0}} \prod_{n=1}^M \frac{\Psi_{\lambda_n}(\Gamma_{n-1})}{\Psi_{\lambda_{n-1}}(\Gamma_{n-1})}. \quad (25)$$

Substituting this expression into Eq. (22), we obtain

$$\langle e^{-\beta W} \rangle = \frac{Z_{\lambda_M}}{Z_{\lambda_0}} \int d\Gamma_0 \cdots d\Gamma_M \Psi_{\lambda_0}(\Gamma_0) \times \prod_{n=1}^M \frac{R_{\lambda_n}(\Gamma_n | \Gamma_{n-1}) \Psi_{\lambda_n}(\Gamma_{n-1})}{\Psi_{\lambda_{n-1}}(\Gamma_{n-1})}. \quad (26)$$

By using the balance condition [Eq. (24)], the integrals can be carried out one by one—starting with $\int d\Gamma_0$, then $\int d\Gamma_1$, and so forth. For example, the integral over Γ_0 is calculated as

$$\int d\Gamma_0 \Psi_{\lambda_0}(\Gamma_0) \frac{R_{\lambda_1}(\Gamma_1 | \Gamma_0) \Psi_{\lambda_1}(\Gamma_0)}{\Psi_{\lambda_0}(\Gamma_0)} = \Psi_{\lambda_1}(\Gamma_1). \quad (27)$$

After carrying out all the integrals, we obtain Jarzynski's equality,

$$\langle e^{-\beta W} \rangle = \frac{Z_{\lambda_M}}{Z_{\lambda_0}} = \frac{Z_{\Lambda}}{Z_0} = e^{-\beta \Delta F}. \quad (28)$$

D. Jarzynski's equality for isobaric–isothermal systems

It is a natural attempt to generalize Jarzynski's equality, which was originally derived for isothermal²⁹ processes (i.e., for the canonical ensemble), to other statistical ensembles. Inspecting the derivation of Jarzynski's equality, one realizes that it is based on the same exponential form shared by $e^{-\beta W}$ and the Boltzmann factor $e^{-\beta H}$. Therefore, for any ensemble described by an exponential weighting factor a similar equality is expected. In particular, if a system is in contact with a heat–volume bath at constant temperature T and pressure P , its equilibrium states are distributed according to the NPT (constant number, pressure, and temperature) ensemble which has the exponential weighting factor $e^{-\beta(H+PV)}$, with V being the volume of the system. In this section we prove the following: *For a transformation between two equilibrium states of a system in contact with a heat–volume bath at temperature T and pressure P , the work W done on the system during the transformation and the Gibbs free energy difference ΔG between the two equilibrium states satisfy*

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta G}. \quad (29)$$

The equality here is the same as the original Jarzynski's equality except that the Helmholtz free energy is replaced by the Gibbs free energy.³⁰ This is a useful result, as experiments and computer simulations in biophysics are often performed at constant temperature and pressure. We will call Eq. (29) the *isobaric–isothermal Jarzynski equality* and the original Jarzynski's equality the *isothermal Jarzynski equality*.

The difference from the isothermal case is that the volume V of the system fluctuates and hence needs to be specified, in addition to the phase Γ , in order to determine a microscopic state of the system. The Hamiltonian $H_{\lambda}(\Gamma, V)$, a function of both Γ and V in general, is assumed to depend on some parameter λ which is controlled externally. For each value of λ , the NPT partition function Y_{λ} , the Gibbs free energy G_{λ} , and the equilibrium probability distribution $\Psi_{\lambda}(\Gamma, V)$ are given as

$$Y_{\lambda} = \int dV \int_{\Gamma} d\Gamma \exp[-\beta H_{\lambda}(\Gamma, V) - \beta PV], \quad (30a)$$

$$G_{\lambda} = -\frac{1}{\beta} \log Y_{\lambda}, \quad (30b)$$

$$\Psi_{\lambda}(\Gamma, V) = \frac{1}{Y_{\lambda}} \exp[-\beta H_{\lambda}(\Gamma, V) - \beta PV], \quad (30c)$$

where $\int_{\Gamma} d\Gamma$ denotes an integral over all possible atomic positions contained in volume V and all possible atomic momenta. Hereafter we denote (Γ, V) by X .

To prove the isobaric–isothermal Jarzynski equality, we follow the approach in Sec. II C. Since it is a rather straightforward generalization, only a sketch of the basic steps of the derivation will be given. The dynamics of the system is de-

scribed by stochastic dynamics, without explicitly accounting for the degrees of freedom of the bath. Again two assumptions are made: the Markov property

$$\partial_t f(X, t) = \mathcal{L}_{\lambda(t)} f(X, t) \quad (31)$$

and the balance condition

$$\mathcal{L}_{\lambda} \Psi_{\lambda}(X) = 0. \quad (32)$$

Based on the same discretization scheme as in Sec. II C, the work done on the system is given as

$$W = \sum_{n=1}^M [H_{\lambda_n}(X_{n-1}) - H_{\lambda_{n-1}}(X_{n-1})]. \quad (33)$$

The average $\langle e^{-\beta W} \rangle$ can be written in terms of transition probabilities,

$$\begin{aligned} \langle e^{-\beta W} \rangle &= \int dX_0 \cdots dX_M \Psi_{\lambda_0}(X_0) \\ &\quad \times \left[\prod_{n=1}^M R_{\lambda_n}(X_n | X_{n-1}) \right] e^{-\beta W}. \end{aligned} \quad (34)$$

The balance condition takes the form

$$\Psi_{\lambda_n}(X_n) = \int dX_{n-1} R_{\lambda_n}(X_n | X_{n-1}) \Psi_{\lambda_n}(X_{n-1}). \quad (35)$$

We can write $e^{-\beta W}$ in terms of the equilibrium probability Ψ ,

$$\begin{aligned} e^{-\beta W} &= \prod_{n=1}^M \frac{\exp[-\beta H_{\lambda_n}(X_{n-1})]}{\exp[-\beta H_{\lambda_{n-1}}(X_{n-1})]} \\ &= \prod_{n=1}^M \frac{\exp[-\beta H_{\lambda_n}(X_{n-1}) - \beta PV]}{\exp[-\beta H_{\lambda_{n-1}}(X_{n-1}) - \beta PV]} \\ &= \frac{Y_{\lambda_M}}{Y_{\lambda_0}} \prod_{n=1}^M \frac{\Psi_{\lambda_n}(X_{n-1})}{\Psi_{\lambda_{n-1}}(X_{n-1})}. \end{aligned} \quad (36)$$

Upon substitution of Eq. (36) into Eq. (34) and completion of the integrals, from $\int dX_0$ up to $\int dX_M$, we obtain the isobaric–isothermal Jarzynski equality,

$$\langle e^{-\beta W} \rangle = \frac{Y_{\lambda_M}}{Y_{\lambda_0}} = e^{-\beta \Delta G}. \quad (37)$$

E. Jarzynski's equality and the second law of thermodynamics

From Jarzynski's equality [Eq. (1)] and Jensen's inequality ($\langle e^x \rangle \geq e^{\langle x \rangle}$) follows

$$\langle W \rangle \geq \Delta F, \quad (38)$$

where the equality sign holds if and only if all the sampled work values are equal, i.e., if and only if the variance of the work W vanishes.

The following is a direct implication of the second law of thermodynamics for isothermal processes (Ref. 31, Sec. 13): *for a transformation between two equilibrium states of a system in contact with a heat bath at a constant temperature,*

the work done on the system during the transformation is not smaller than the Helmholtz free energy difference between the two equilibrium states, namely,

$$W \geq \Delta F, \tag{39}$$

where the equality sign holds if and only if the transformation is reversible. Jarzynski's equality tells us that this is true on average. For a single realization Eq. (39) might well be violated. The chance of violation, however, is very small. Let $P(W)$ be the probability distribution for the work W . Then the probability for observing a violation by the amount of D or larger is $\int_{-\infty}^{\Delta F - D} dWP(W)$. As shown in Ref. 32, from the inequality chain

$$\begin{aligned} e^{-\beta \Delta F} &= \int_{-\infty}^{\infty} dWP(W) e^{-\beta W} \\ &\geq \int_{-\infty}^{\Delta F - D} dWP(W) e^{-\beta W} \\ &\geq e^{-\beta(\Delta F - D)} \int_{-\infty}^{\Delta F - D} dWP(W) \end{aligned} \tag{40}$$

it follows

$$\int_{-\infty}^{\Delta F - D} dWP(W) \leq e^{-\beta D}. \tag{41}$$

If D is a macroscopic quantity, $e^{-\beta D}$ is extremely small; macroscopic violations of the second law are prohibited.

The equality sign in the second law [Eq. (39)] holds when the transformation is reversible. On the other hand, the equality sign in Eq. (38) holds when the work distribution has a vanishing variance. (When the variance is zero, all the sampled work values are the same and therefore the equality sign in Jensen's inequality holds.) Generally, the variance of work decreases as the transformation slows down, and reaches zero in the reversible limit. This can be established using the discretized framework of Sec. II C. When the transformation is sufficiently slow, the total M steps (Fig. 1) can be divided into I intervals, each containing S steps ($M = IS$) such that (i) $S\delta t$ is much longer than the correlation time of the stochastic dynamics and (ii) the Hamiltonian changes negligibly over S steps. Let ϵ be the increment of the parameter λ over S steps: $\epsilon = S\delta\lambda$. The total work W can be written as the sum of the work done in each interval,

$$\begin{aligned} W &= \sum_{i=1}^I w_i, \\ w_i &= \sum_{n=S(i-1)+1}^{Si} [H_{\lambda_n}(\Gamma_{n-1}) - H_{\lambda_{n-1}}(\Gamma_{n-1})]. \end{aligned} \tag{42}$$

As $\epsilon \rightarrow 0$, which is approached as the transformation slows down, w_i satisfies

$$w_i \sim \epsilon, \quad \text{var}(w_i) \sim \epsilon^2. \tag{43}$$

Since w_i 's at different i 's are uncorrelated ($S\delta t$ is much longer than the correlation time), we find

$$\text{var}(W) = \sum_{i=1}^I \text{var}(w_i) \sim \sum_{i=1}^I \epsilon^2 \sim \epsilon, \tag{44}$$

which vanishes as $\epsilon \rightarrow 0$.

In summary, for isothermal processes the second law of thermodynamics is derived from Jarzynski's equality. A similar argument applies to isobaric–isothermal processes, with the Gibbs free energy playing the role of the Helmholtz free energy. Another class of processes studied in thermodynamics are those in which a *thermally isolated* system undergoes an external operation (by changing some parameter). If the process is reversible, the entropy of the system remains constant; if irreversible, the entropy increases. It is an interesting question whether one can find for such processes an equality similar to Jarzynski's.

F. Jarzynski's equality and computer simulations

Computer simulations cannot explicitly include baths of infinite size (or much larger than the systems of interest). However, it is possible to simulate the *effect* of baths, and various algorithms have been developed in this regard.^{33,34} Jarzynski's equality is obeyed by most of those algorithms. This is not surprising because (i) most simulation algorithms are history-independent and hence represent Markov processes; (ii) the balance condition is a minimal effect of baths, and accordingly any operative algorithm that simulates a bath is expected to satisfy the balance condition.

Monte Carlo and molecular dynamics are two major variants of molecular simulation methods. Monte Carlo is obviously a Markov process since it is history-independent. And Monte Carlo simulations, either isothermal or isobaric–isothermal, satisfy the balance condition because they are implemented based on detailed balance which is an even stronger condition. As such, Monte Carlo satisfies the two conditions for Jarzynski's equality.¹¹

One way to incorporate the effect of baths in molecular dynamics simulations is to include additional terms (usually friction and random noise) in the equation of motion in such a way that the resulting trajectories sample the appropriate statistical ensemble. The Langevin dynamics method for isothermal simulations and the Langevin piston method²⁸ for isobaric–isothermal simulations belong to this category. The resulting equation of motion is a stochastic differential equation due to the random noise term and can be converted to a Fokker–Planck equation which is of the form of Eq. (12) (Ref. 35, Sec. 4.3.4). Then one only needs to check the balance condition, i.e., whether the equilibrium distribution is stationary under the Fokker–Planck equation. Jarzynski¹¹ did exactly this for Langevin dynamics and confirmed that it satisfies the balance condition. We will show in Sec. II G that the Langevin piston method, too, satisfies the balance condition.

Another way to incorporate the effect of baths in molecular dynamics simulations is to introduce additional degrees of freedom while retaining the deterministic nature of the dynamics.³⁶ The resulting trajectories, when projected onto the space of the original degrees of freedom, are supposed to sample the appropriate ensemble. Nosé–Hoover thermostat^{37,38} for isothermal molecular dynamics is a typical example. Even when the dynamics is deterministic, an equation of the form of Eq. (12) can still be written if one considers ensembles; an example is provided by the Liouville

equation in classical mechanics (Ref. 39, Sec. 9.8). Again, one only needs to check whether the equilibrium distribution is stationary under the resulting Liouville-type equation. Jarzynski^{10,11} confirmed that the Nosé–Hoover thermostat indeed satisfies this.

G. The Langevin piston method for isobaric–isothermal molecular dynamics

The Langevin piston method²⁸ is widely used for isobaric–isothermal molecular dynamics simulations. With the Hamiltonian

$$H_\lambda(\mathbf{r}, \mathbf{p}) = \sum_{i=1}^{3N} \frac{p_i^2}{2m_i} + U_\lambda(\mathbf{r}), \quad (45)$$

the Langevin piston method involves the following stochastic differential equations:

$$\dot{r}_i = \frac{p_i}{m_i} + \frac{b}{3wV} r_i, \quad (46a)$$

$$\dot{p}_i = -\frac{\partial}{\partial r_i} U_{\lambda(t)}(\mathbf{r}) - \frac{b}{3wV} p_i - \frac{\gamma_i}{m_i} p_i + \sigma_i \eta_i, \quad (46b)$$

$$\dot{V} = \frac{b}{w}, \quad (46c)$$

$$\dot{b} = B(\mathbf{r}, \mathbf{p}, V, t) - P - \frac{\alpha}{w} b + \rho \mu. \quad (46d)$$

Here r_i , p_i , and m_i are atomic positions, momenta, and masses, respectively; b and w are the effective momentum and mass, respectively, associated with the volume V . At each instant, pressure is estimated through the virial equation (see Appendix)

$$B(\mathbf{r}, \mathbf{p}, V, t) = \frac{1}{3V} \sum_{i=1}^{3N} \left[\frac{p_i^2}{m_i} - r_i \frac{\partial}{\partial r_i} U_{\lambda(t)}(\mathbf{r}) \right] \quad (47)$$

and is controlled toward P , the imposed pressure. For temperature control, white noise variables η_i and μ are used,

$$\begin{aligned} \langle \eta_i(t) \eta_j(t') \rangle &= \delta_{ij} \delta(t-t'), \\ \langle \mu(t) \mu(t') \rangle &= \delta(t-t'), \quad \langle \eta_i(t) \mu(t') \rangle = 0. \end{aligned} \quad (48)$$

The parameters γ_i , σ_i , α , and ρ that represent the strengths of friction and noise are chosen such that they obey the fluctuation–dissipation relations

$$\sigma_i^2 = 2\gamma_i k_B T, \quad \rho^2 = 2\alpha k_B T. \quad (49)$$

In the original formulation of the Langevin piston method,²⁸ only the volume degree of freedom is used for temperature control; in other words, γ_i and σ_i are set to zero. However, the additional temperature control leads to faster relaxation of energy. The following arguments apply to either case.

The stochastic differential equations [Eq. (46)] govern the time evolution of the microscopic state $(\mathbf{r}, \mathbf{p}, V, b)$. Notice that b , the momentum associated with the volume, is included. From these stochastic differential equations follows the Fokker–Planck equation for the probability distribution $f(\mathbf{r}, \mathbf{p}, V, b, t)$,

$$\begin{aligned} \frac{\partial}{\partial t} f &= - \sum_{i=1}^{3N} \frac{\partial}{\partial r_i} \left[\left(\frac{p_i}{m_i} + \frac{b}{3wV} r_i \right) f \right] \\ &+ \sum_{i=1}^{3N} \frac{\partial}{\partial p_i} \left[\left(\frac{\partial}{\partial r_i} U_{\lambda(t)}(\mathbf{r}) + \frac{b}{3wV} p_i + \frac{\gamma_i}{m_i} p_i \right) f \right] \\ &- \frac{\partial}{\partial V} \left(\frac{b}{w} f \right) - \frac{\partial}{\partial b} \left[\left(B(\mathbf{r}, \mathbf{p}, V, t) - P - \frac{\alpha}{w} b \right) f \right] \\ &+ \sum_{i=1}^{3N} \frac{\sigma_i^2}{2} \frac{\partial^2}{\partial p_i^2} f + \frac{\rho^2}{2} \frac{\partial^2}{\partial b^2} f. \end{aligned} \quad (50)$$

After some tedious but straightforward algebra, it can be shown that the equilibrium distribution $\Psi_\lambda(\mathbf{r}, \mathbf{p}, V, b)$ given as

$$\Psi_\lambda(\mathbf{r}, \mathbf{p}, V, b) = \frac{1}{Y_\lambda} \exp[-\beta H_\lambda(\mathbf{r}, \mathbf{p}) - \beta P V - \beta b^2/2w], \quad (51a)$$

$$Y_\lambda = \int db \int dV \int d\mathbf{p} \int_V d\mathbf{r} \exp[-\beta H_\lambda(\mathbf{r}, \mathbf{p}) - \beta P V - \beta b^2/2w] \quad (51b)$$

is a stationary solution of the Fokker–Planck equation. Therefore we conclude that if the initial states are sampled from $\Psi_0(\mathbf{r}, \mathbf{p}, V, b)$, the isobaric–isothermal Jarzynski equality [Eq. (29)] holds. Notice that the Gibbs free energy $G_\lambda = -k_B T \log Y_\lambda$ in this case has an additional term due to the additional degree of freedom, b [compare Eqs. (30a) and (51b)]. This additional term, however, is canceled out in the difference ΔG .

On an additional note, the stationarity of the equilibrium distribution turns out to be sensitive to the form of the virial equation. Often, p_i^2/m_i in the virial equation [Eq. (47)] is replaced by its thermal average, $k_B T$, and the following form is used:³³

$$B(\mathbf{r}, \mathbf{p}, V, t) = \frac{Nk_B T}{V} - \frac{1}{3V} \sum_{i=1}^{3N} r_i \frac{\partial}{\partial r_i} U_{\lambda(t)}(\mathbf{r}). \quad (52)$$

With this alternative form, however, the equilibrium distribution is no longer stationary.

III. CALCULATING POTENTIALS OF MEAN FORCE FROM STEERED MOLECULAR DYNAMICS SIMULATIONS

In this section we present a method for calculating PMFs from SMD simulations and discuss related issues. The method is based on Jarzynski's equality and the choice of a large spring constant for the guiding potential. Since the exponential average appearing in Jarzynski's equality is difficult to evaluate, the cumulant expansion is employed as an approximation. We discuss the possibility that SMD simulations through the use of stiff springs can be made to conform to Gaussian work distributions for which the cumulant expansion for PMFs can be safely terminated at second order.

A. PMF, SMD, and Jarzynski's equality

Consider a classical mechanical system of N particles in contact with a heat bath at constant temperature T . A microscopic state is specified by $3N$ -dimensional position \mathbf{r} and momentum \mathbf{p} . Suppose that we have identified a reaction coordinate $\xi(\mathbf{r})$. The PMF $\Phi(\xi)$ along ξ is defined by

$$\exp[-\beta\Phi(\xi')] = \int d\mathbf{r}d\mathbf{p} \delta(\xi(\mathbf{r}) - \xi') \exp[-\beta H(\mathbf{r}, \mathbf{p})], \quad (53)$$

where $\beta = 1/k_B T$ is the inverse temperature and H is the Hamiltonian. The PMF $\Phi(\xi)$ is the Helmholtz free energy profile along the reaction coordinate ξ ; the probability of observing the reaction coordinate at ξ is proportional to $\exp[-\beta\Phi(\xi)]$. If the system is in contact with a heat-volume bath at constant temperature T and pressure P , the corresponding PMF is defined by

$$\exp[-\beta\Phi(\xi')] = \int dV \int d\mathbf{p} \int_V d\mathbf{r} \delta(\xi(\mathbf{r}) - \xi') \times \exp[-\beta H(\mathbf{r}, \mathbf{p}, V) - \beta P V], \quad (54)$$

where $\int_V d\mathbf{r}$ denotes an integral over positions contained in the volume V . In this case the PMF is the Gibbs free energy profile along the reaction coordinate. For the sake of simplicity, we will work within the isothermal framework. The generalization to the isobaric-isothermal framework is straightforward.

SMD is an efficient way to explore the system along the reaction coordinate. In a SMD simulation a guiding potential

$$h_\lambda(\mathbf{r}) = \frac{k}{2} [\xi(\mathbf{r}) - \lambda]^2 \quad (55)$$

is added to the original Hamiltonian H . We write the total Hamiltonian as

$$\tilde{H}_\lambda(\mathbf{r}, \mathbf{p}) = H(\mathbf{r}, \mathbf{p}) + h_\lambda(\mathbf{r}). \quad (56)$$

The parameter λ is changed typically with a constant velocity,

$$\lambda(t) = \lambda(0) + vt, \quad (57)$$

covering the relevant region of ξ . Atomic force microscopy experiments can be accounted for by the same procedure.⁴⁰

Applying Jarzynski's equality to the \tilde{H} -system, we obtain

$$F_{\lambda(\tau)} - F_{\lambda(0)} = -\frac{1}{\beta} \log \langle \exp[-\beta W(\tau)] \rangle. \quad (58)$$

Here F_λ is the Helmholtz free energy of the \tilde{H} -system,

$$\exp(-\beta F_\lambda) = \int d\mathbf{r}d\mathbf{p} \exp[-\beta \tilde{H}_\lambda(\mathbf{r}, \mathbf{p})], \quad (59)$$

and $W(\tau)$ is the work done on the \tilde{H} -system during the time interval between zero and τ , calculated for each trajectory $(\mathbf{r}(t), \mathbf{p}(t))$ as

$$W(\tau) = \int_0^\tau dt \left[\frac{\partial}{\partial t} \tilde{H}_{\lambda(t)}(\mathbf{r}, \mathbf{p}) \right]_{(\mathbf{r}, \mathbf{p}) = (\mathbf{r}(t), \mathbf{p}(t))}. \quad (60)$$

The average $\langle \cdot \rangle$ in Eq. (58) is taken over the ensemble of trajectories the initial states $(\mathbf{r}(0), \mathbf{p}(0))$ of which are sampled from the canonical ensemble corresponding to $\tilde{H}_{\lambda(0)}(\mathbf{r}(0), \mathbf{p}(0))$.

In the following section we use the so-called stiff-spring approximation in order to extract $\Phi(\xi)$, the PMF of the original H -system, from F_λ , the free energy of the \tilde{H} -system.

B. The stiff-spring approximation

At a certain instant t , the value of the parameter λ is fixed at $\lambda(t) = \lambda(0) + vt$. The reaction coordinate $\xi(\mathbf{r}(t))$, on the other hand, may take any value, though the guiding potential [Eq. (55)] holds it near $\lambda(t)$. The idea of the stiff-spring approximation is to minimize the fluctuation of the reaction coordinate among different trajectories by choosing a sufficiently large spring constant k for the guiding potential.

The free energy F_λ can be expressed in terms of the PMF $\Phi(\xi)$ as follows:

$$\begin{aligned} \exp(-\beta F_\lambda) &= \int d\mathbf{r}d\mathbf{p} \exp\left\{-\beta H(\mathbf{r}, \mathbf{p}) - \frac{\beta k}{2} [\xi(\mathbf{r}) - \lambda]^2\right\} \\ &= \int d\mathbf{r}d\mathbf{p} \int d\xi' \delta(\xi(\mathbf{r}) - \xi') \\ &\quad \times \exp\left\{-\beta H(\mathbf{r}, \mathbf{p}) - \frac{\beta k}{2} [\xi(\mathbf{r}) - \lambda]^2\right\} \\ &= \int d\xi \exp\left[-\beta\Phi(\xi) - \frac{\beta k}{2} (\xi - \lambda)^2\right]. \quad (61) \end{aligned}$$

When k is large, most of the contribution to this integral comes from the region around $\xi = \lambda$. Thus we take the Taylor series of $\exp[-\beta\Phi(\xi)]$ about λ ,

$$\begin{aligned} \exp[-\beta\Phi(\xi)] &= \exp[-\beta\Phi(\lambda)] \left\{ 1 - \beta \frac{\partial\Phi(\lambda)}{\partial\lambda} (\xi - \lambda) \right. \\ &\quad - \frac{\beta}{2} \left[\frac{\partial^2\Phi(\lambda)}{\partial\lambda^2} - \beta \left(\frac{\partial\Phi(\lambda)}{\partial\lambda} \right)^2 \right] (\xi - \lambda)^2 \\ &\quad \left. + \dots \right\}, \quad (62) \end{aligned}$$

and then obtain the expansion of Eq. (61) about $k = \infty$ by calculating the integral for each term,

$$\begin{aligned} \exp(-\beta F_\lambda) &= \exp[-\beta\Phi(\lambda)] \sqrt{\frac{2\pi}{\beta k}} \left\{ 1 - \frac{1}{2k} \left[\frac{\partial^2\Phi(\lambda)}{\partial\lambda^2} \right. \right. \\ &\quad \left. \left. - \beta \left(\frac{\partial\Phi(\lambda)}{\partial\lambda} \right)^2 \right] + O(1/k^2) \right\}. \quad (63) \end{aligned}$$

Upon taking the logarithm and dropping the irrelevant terms that are independent of λ , we find

$$F_\lambda = \Phi(\lambda) - \frac{1}{2k} \left(\frac{\partial\Phi(\lambda)}{\partial\lambda} \right)^2 + \frac{1}{2\beta k} \frac{\partial^2\Phi(\lambda)}{\partial\lambda^2} + O(1/k^2), \quad (64)$$

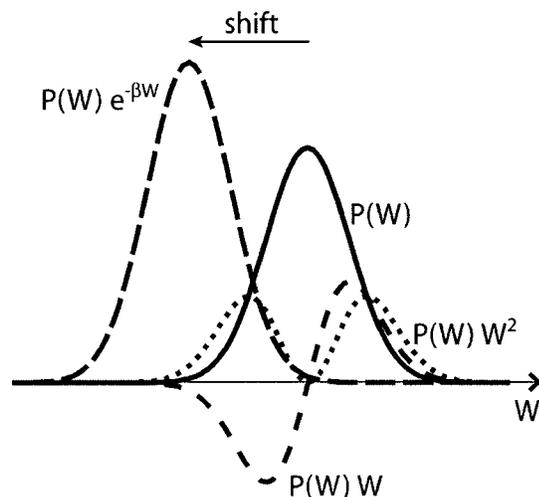


FIG. 2. Difficulty of estimating the exponential average. Typically, the peak of $P(W)e^{-\beta W}$ is shifted from that of the work distribution $P(W)$. This makes $\langle e^{-\beta W} \rangle$ difficult to estimate. On the other hand, $\langle W \rangle$ and $\langle W^2 \rangle$ are easier to estimate because $P(W)W$ and $P(W)W^2$ are centered around the peak of $P(W)$.

which is inverted to

$$\Phi(\lambda) = F_\lambda + \frac{1}{2k} \left(\frac{\partial F_\lambda}{\partial \lambda} \right)^2 - \frac{1}{2\beta k} \frac{\partial^2 F_\lambda}{\partial \lambda^2} + O(1/k^2). \quad (65)$$

Higher order terms can be obtained in a similar way.

In a practical application, one chooses the spring constant k large enough that the fluctuation of the reaction coordinate among different trajectories is minimized, or smaller than the resolution one seeks. A number of trajectories are generated by repeating the SMD simulation with initial conditions sampled from the initial canonical ensemble, and the work $W(\tau)$ is calculated as a function of the final time τ for each trajectory [Eq. (60)]. The free energy F_λ is then calculated as a function of λ by using Eq. (58) or the cumulant expansion [Eq. (67)] which will be explained shortly. The PMF Φ is obtained from Eq. (65) up to a certain order in $1/k$; the next order can be used for checking the validity of the stiff-spring approximation. In the simplest case the PMF is calculated from the leading order, $\Phi(\lambda) = F_\lambda$, which is justified if the first order term turns out to be small.

C. Cumulant expansion

The major difficulty in the use of Jarzynski's equality is that the exponential average $\langle e^{-\beta W} \rangle$ is dominated by small work values that arise only rarely. An accurate estimate of PMF hence requires proper sampling of those rare trajectories that result in small work values. This point is illustrated in Fig. 2. Let $P(W)$ be the probability distribution of the work, which is typically of a bell shape. Then $P(W)e^{-\beta W}$ is another bell-shaped function, but with its peak shifted toward the left from that of $P(W)$. Most work values are sampled around the peak of $P(W)$, whereas the exponential average $\int dW P(W)e^{-\beta W}$ cannot be estimated accurately without properly sampling the region around the peak of $P(W)e^{-\beta W}$. For example, assume that $P(W)$ is a Gaussian with a width (defined as the standard deviation) of σ . Then

$P(W)e^{-\beta W}$ is another (unnormalized) Gaussian with the same width, but with its peak shifted toward the left by $\beta\sigma^2$. When the shift is much larger than the width, there is little overlap between $P(W)$ and $P(W)e^{-\beta W}$, which makes the estimate of the exponential average impractical; it is practical only when the shift-to-width ratio $\beta\sigma$ is not too large, namely when the work fluctuation σ is not much larger than the temperature $k_B T$.

Because of the difficulty in estimating the exponential average, the cumulant expansion is often employed.^{10,19,25} The logarithm of an exponential average can be expanded in terms of cumulants,

$$\log \langle e^x \rangle = \langle x \rangle + \frac{1}{2} (\langle x^2 \rangle - \langle x \rangle^2) + \dots, \quad (66)$$

where the first and second cumulants are shown. Marcinkiewicz's theorem⁴¹ states that either (i) all but the first two cumulants vanish or (ii) there are an infinite number of non-vanishing cumulants. The first case happens if and only if the variable x is sampled from a Gaussian distribution. Using this expansion in Eq. (58), we obtain the cumulant expansion formula for the free energy,

$$F_{\lambda(\tau)} - F_{\lambda(0)} = \langle W(\tau) \rangle - \frac{\beta}{2} (\langle W(\tau)^2 \rangle - \langle W(\tau) \rangle^2) + \dots. \quad (67)$$

An approximate formula is obtained by terminating the series at a certain order. In fact, the second order formula is identical with the near-equilibrium formula^{42,43} predating Jarzynski's equality.

When we use an approximate formula based on the cumulant expansion, two kinds of error are involved: the error due to the truncation of higher order terms and the error due to insufficient sampling. If we use the exact formula [Eq. (58)], we will have no truncation error, but will have possibly a big sampling error because of the difficulty in estimating the exponential average. On the other hand, low order cumulants are relatively easier to estimate from limited sampling. Figure 2 illustrates this showing that the curves $P(W)W$ and $P(W)W^2$ are centered around the peak of $P(W)$ while the curve $P(W)e^{-\beta W}$ is shifted away from it. Thus, for limited sampling an approximate formula may work better than the exact formula. Especially the second order cumulant expansion formula has proven to be effective in SMD simulations.^{23,25}

The most fortunate case arises when the work distribution is Gaussian, for which the second order formula can be used without the penalty of a truncation error. For slow processes, the work distribution is expected to be Gaussian as suggested by the near-equilibrium formula.^{42,43} For processes of arbitrary speeds, in general, the work distribution may not be Gaussian. In the following we argue that a SMD simulation performed with a stiff spring leads to a Gaussian work distribution regardless of the speed of the process; this may explain the success of the second order formula in previous applications.

D. The Gaussian nature of the work distribution

Consider a SMD simulation performed along a reaction coordinate ξ with a moving guiding potential $(k/2)(\xi$

$-vt)^2$. [For simplicity we set $\lambda(0)=0$ in Eq. (57), which can be done by shifting the origin of the reaction coordinate.] Let us assume that the motion along the reaction coordinate can be described by the overdamped Langevin equation, which is frequently used for modeling biomolecular processes,

$$\frac{d\xi}{dt} = -\beta D(\xi) \frac{\partial}{\partial \xi} U(\xi, t) + \sqrt{2D(\xi)} \eta, \quad (68)$$

with a white noise variable η ,

$$\langle \eta(t) \eta(t') \rangle = \delta(t-t'). \quad (69)$$

The diffusion coefficient D , in general, is ξ -dependent. The potential $U(\xi, t)$ is the sum of the PMF and the moving guiding potential,

$$U(\xi, t) = \Phi(\xi) + \frac{k}{2} (\xi - vt)^2. \quad (70)$$

The initial condition $\xi(0)$ is sampled from the equilibrium distribution corresponding to the initial potential $U(\xi, 0)$ at temperature T . If both D and Φ are constant, Eq. (68) describes the diffusion on a moving harmonic potential, for which the work distribution is Gaussian.⁴⁴ In the following we show that if the spring constant k is sufficiently large, the dynamics (after a change of variables) is governed by essentially the same equation as the diffusion on a moving harmonic potential and therefore the work distribution is Gaussian.

Let us assume that k is chosen so large that the reaction coordinate ξ is always close to the center of the guiding potential, vt . Then the potential U can be approximated as

$$U(\xi, t) \approx \Phi(vt) + \Phi'(vt)(\xi - vt) + \frac{k}{2} (\xi - vt)^2, \quad (71)$$

and the overdamped Langevin equation [Eq. (68)] as

$$\frac{d\xi}{dt} \approx -\beta D(vt) [k(\xi - vt) + \Phi'(vt)] + \sqrt{2D(vt)} \eta. \quad (72)$$

The external work done between time zero and t is given by Eq. (60), with \tilde{H} replaced by $U(\xi, t)$ in Eq. (70),

$$W(t) = -vk \int_0^t dt' [\xi(t') - vt']. \quad (73)$$

Taking the derivative, we obtain

$$\frac{dW}{dt} = -vk(\xi - vt). \quad (74)$$

Equations (72) and (74) constitute a system of stochastic differential equations.

The following change of variables, $(\xi, W) \rightarrow (\zeta, \Omega)$, proves to be useful:

$$\zeta = \xi - vt + \frac{1}{k} \Phi'(vt), \quad (75a)$$

$$\Omega = W - [\Phi(vt) - \Phi(0)]. \quad (75b)$$

The new variable ζ is the deviation of the reaction coordinate from the instantaneous minimum of the potential U , and Ω is

the irreversible (dissipative) work. (The reversible work is the same as the change in the PMF Φ .) This change of variables leads to, using Eqs. (72) and (74),

$$\begin{aligned} \frac{d\zeta}{dt} &= \frac{d\xi}{dt} - v + \frac{v}{k} \Phi''(vt) \\ &\approx \frac{d\xi}{dt} - v = -\beta k D(vt) \zeta - v + \sqrt{2D(vt)} \eta, \end{aligned} \quad (76a)$$

$$\frac{d\Omega}{dt} = \frac{dW}{dt} - v \Phi'(vt) = -vk\zeta. \quad (76b)$$

From these stochastic differential equations follows the Fokker-Planck equation for the probability distribution $P(\zeta, \Omega, t)$,

$$\begin{aligned} \frac{\partial P}{\partial t} = \mathcal{L}P = &\left\{ \beta k D(vt) + [v + \beta k D(vt) \zeta] \frac{\partial}{\partial \zeta} \right. \\ &\left. + D(vt) \frac{\partial^2}{\partial \zeta^2} + vk\zeta \frac{\partial}{\partial \Omega} \right\} P. \end{aligned} \quad (77)$$

The adjoint of the operator \mathcal{L} is

$$\mathcal{L}^\dagger = -[v + \beta k D(vt) \zeta] \frac{\partial}{\partial \zeta} + D(vt) \frac{\partial^2}{\partial \zeta^2} - vk\zeta \frac{\partial}{\partial \Omega}. \quad (78)$$

The initial distribution

$$P(\zeta, \Omega, 0) = \sqrt{\frac{\beta k}{2\pi}} \exp\left(-\frac{\beta k}{2} \zeta^2\right) \delta(\Omega), \quad (79)$$

which is Gaussian, serves as the initial condition for the Fokker-Planck equation.

It proves to be more effective to work with the *cumulant generating function*

$$Q(s, u, t) = \log \int_{-\infty}^{\infty} d\zeta \int_{-\infty}^{\infty} d\Omega \exp(is\zeta + iu\Omega) P(\zeta, \Omega, t) \quad (80)$$

than dealing with the probability distribution P directly. Notice that P is completely determined by Q through the inverse Fourier transform

$$\begin{aligned} P(\zeta, \Omega, t) &= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} ds \int_{-\infty}^{\infty} du \\ &\times \exp(-is\zeta - iu\Omega) \exp Q(s, u, t). \end{aligned} \quad (81)$$

When Q is expanded as a power series in (s, u) , the coefficients give cumulants (Ref. 35, Sec. 2.7),

$$\begin{aligned} Q(s, u, t) &= i\langle \zeta(t) \rangle_s + i\langle \Omega(t) \rangle_u - \langle \zeta(t) \Omega(t) \rangle_{cs} u \\ &\quad - \frac{1}{2} \langle \zeta(t)^2 \rangle_{cs} s^2 - \frac{1}{2} \langle \Omega(t)^2 \rangle_{cu} u^2 + \dots \end{aligned} \quad (82)$$

Cumulants can be expressed in terms of moments

$$\langle \zeta \Omega \rangle_c = \langle \zeta \Omega \rangle - \langle \zeta \rangle \langle \Omega \rangle, \quad \langle \zeta^2 \rangle_c = \langle \zeta^2 \rangle - \langle \zeta \rangle^2, \quad (83)$$

and so on. By Marcinkiewicz's theorem,⁴¹ the degree of the power series [Eq. (82)] is either two (for Gaussian distributions) or infinity.

Using Eqs. (77) and (80), we obtain a differential equation governing the time evolution of the cumulant generating function,

$$\begin{aligned} \frac{\partial Q}{\partial t} &= e^{-Q} \int d\zeta d\Omega \exp(is\zeta + iu\Omega) \frac{\partial P}{\partial t} \\ &= e^{-Q} \int d\zeta d\Omega \exp(is\zeta + iu\Omega) \mathcal{L}P \\ &= e^{-Q} \int d\zeta d\Omega P \mathcal{L}^\dagger \exp(is\zeta + iu\Omega) \\ &= e^{-Q} \int d\zeta d\Omega P \{ -ivs - D(vt)s^2 \\ &\quad - [i\beta k D(vt)s + ivku] \zeta \} \exp(is\zeta + iu\Omega) \\ &= e^{-Q} \left\{ -ivs - D(vt)s^2 - [\beta k D(vt)s + vku] \frac{\partial}{\partial s} \right\} e^Q \\ &= -ivs - D(vt)s^2 - [\beta k D(vt)s + vku] \frac{\partial Q}{\partial s}. \quad (84) \end{aligned}$$

An important property of this differential equation is that if Q at some instant happens to be a second degree polynomial, then $\partial Q/\partial t$ is also a second degree polynomial and Q at any later (or earlier) instant remains to be a second degree polynomial. Therefore, *once P is Gaussian, it is always Gaussian.*

The general Gaussian distribution for two variables (ζ, Ω) can be written in terms of a positive definite correlation matrix \mathbb{C} ,

$$P = \frac{1}{2\pi\sqrt{|\mathbb{C}|}} \exp\left(-\frac{1}{2} Z^T \mathbb{C}^{-1} Z\right), \quad (85a)$$

$$Z = \begin{pmatrix} \zeta - \langle \zeta \rangle \\ \Omega - \langle \Omega \rangle \end{pmatrix}, \quad \mathbb{C} = \begin{pmatrix} \langle \zeta^2 \rangle_c & \langle \zeta \Omega \rangle_c \\ \langle \zeta \Omega \rangle_c & \langle \Omega^2 \rangle_c \end{pmatrix}. \quad (85b)$$

Integrating out ζ gives the probability distribution for Ω ,

$$\int d\zeta P = \frac{1}{\sqrt{2\pi\langle \Omega^2 \rangle_c}} \exp\left[-\frac{(\Omega - \langle \Omega \rangle)^2}{2\langle \Omega^2 \rangle_c}\right], \quad (86)$$

which is Gaussian. The probability for the total work W is also Gaussian because W is linearly related to Ω [Eq. (75b)].

In summary, under the assumption that the overdamped Langevin equation is a good approximation, SMD simulations with stiff springs result in Gaussian work distributions, for which the second order formula of Jarzynski's equality can be used without any truncation error. The idea of using a stiff spring was originally motivated by the need to extract a PMF as a function of a reaction coordinate from a free energy as a function of an external parameter.^{23,25} The use of a stiff spring seems to have another important advantage, namely keeping the work distribution Gaussian.

E. Time evolution of cumulants

The Gaussian distribution P [Eq. (85)] is completely determined by the cumulants, $\langle \zeta \rangle$, $\langle \Omega \rangle$, $\langle \zeta \Omega \rangle_c$, $\langle \zeta^2 \rangle_c$, and

$\langle \Omega^2 \rangle_c$. Therefore, the time evolution of P is determined by the time evolution of the cumulants.⁴⁵ Substituting the cumulant generating function

$$Q = i\langle \zeta \rangle s + i\langle \Omega \rangle u - \langle \zeta \Omega \rangle_c s u - \frac{1}{2} \langle \zeta^2 \rangle_c s^2 - \frac{1}{2} \langle \Omega^2 \rangle_c u^2 \quad (87)$$

into Eq. (84) we find the differential equations governing the time evolution of the cumulants,

$$\frac{d\langle \zeta \rangle}{dt} = -\beta k D(vt) \langle \zeta \rangle - v, \quad (88a)$$

$$\frac{d\langle \Omega \rangle}{dt} = -v k \langle \zeta \rangle, \quad (88b)$$

$$\frac{d\langle \zeta \Omega \rangle_c}{dt} = -v k \langle \zeta^2 \rangle_c - \beta k D(vt) \langle \zeta \Omega \rangle_c, \quad (88c)$$

$$\frac{d\langle \zeta^2 \rangle_c}{dt} = -2\beta k D(vt) \langle \zeta^2 \rangle_c + 2D(vt), \quad (88d)$$

$$\frac{d\langle \Omega^2 \rangle_c}{dt} = -2v k \langle \zeta \Omega \rangle_c. \quad (88e)$$

The accompanying initial condition is obtained from Eq. (79),

$$\begin{aligned} \langle \zeta(0) \rangle &= 0, \quad \langle \Omega(0) \rangle = 0, \quad \langle \zeta(0) \Omega(0) \rangle_c = 0, \\ \langle \zeta(0)^2 \rangle_c &= \frac{1}{\beta k}, \quad \langle \Omega(0)^2 \rangle_c = 0. \end{aligned} \quad (89)$$

Equation (88) is a system of first-order linear ordinary differential equations, and the general solution can be easily written in terms of integrations. However, here we seek simpler approximate solutions. In solving Eq. (88), $\langle \zeta \rangle$, $\langle \zeta \Omega \rangle_c$, and $\langle \zeta^2 \rangle_c$ will feature relaxations (exponential decays) with the time scale of $1/\beta k D$. We assume that these relaxations are much faster than the change in the diffusion coefficient $D(vt)$. In other words, we assume

$$\frac{v}{\beta k D} \ll l, \quad (90)$$

where l is some characteristic length scale over which the diffusion coefficient changes considerably. This assumption, which is likely to be valid because we are using stiff springs, can be checked once the diffusion coefficient is estimated. Under this assumption, we neglect the relaxations and find an approximate solution to Eq. (88),

$$\langle \zeta(t) \rangle = -\frac{v}{\beta k D(vt)}, \quad (91a)$$

$$\langle \Omega(t) \rangle = \int_0^t dt' \frac{v^2}{\beta D(vt')}, \quad (91b)$$

$$\langle \zeta(t) \Omega(t) \rangle_c = -\frac{v}{\beta^2 k D(vt)}, \quad (91c)$$

$$\langle \zeta(t)^2 \rangle_c = \frac{1}{\beta k}, \quad (91d)$$

$$\langle \Omega(t)^2 \rangle_c = \int_0^t dt' \frac{2v^2}{\beta^2 D(vt')} \quad (91e)$$

Let us explore implications of this solution. For this purpose, we rewrite Eq. (91) into the following equivalent equations. (In doing so, we return to the original variables, ξ and W .)

$$\langle W(t) \rangle = -vk \int_0^t dt' [\langle \xi(t') \rangle - vt'], \quad (92a)$$

$$\langle W(t) \rangle - \frac{\beta}{2} \langle W(t)^2 \rangle_c = \Phi(vt) - \Phi(0), \quad (92b)$$

$$\langle W(t)^2 \rangle_c = -2vk \int_0^t dt' \langle \xi(t') W(t') \rangle_c, \quad (92c)$$

$$\langle \xi(t)^2 \rangle_c = \frac{1}{\beta k}, \quad (92d)$$

$$D(vt) = \frac{2v^2}{\beta^2} \left(\frac{d\langle W^2 \rangle_c}{dt} \right)^{-1}. \quad (92e)$$

Equation (92a) follows directly from the definition of the work W [Eq. (73)]; it does not depend on the assumption of the overdamped Langevin equation [Eq. (68)]. Equation (92c) also follows directly from the definition of W ,

$$\begin{aligned} \langle W(t)^2 \rangle_c &= v^2 k^2 \int_0^t dt' \int_0^t dt'' \langle [\xi(t') - vt'] [\xi(t'') - vt''] \rangle_c \\ &= 2v^2 k^2 \int_0^t dt' \int_0^{t'} dt'' \langle [\xi(t') - vt'] [\xi(t'') - vt''] \rangle_c \\ &= -2vk \int_0^t dt' \langle \xi(t') W(t') \rangle_c. \end{aligned} \quad (93)$$

Equation (92b) is nothing but the second order formula of Jarzynski's equality. Equations (92d) and (92e) are consequences of the overdamped Langevin equation and the fast-relaxation condition [Eq. (90)]. Equation (92e), which is a rearrangement of Eq. (91e), can be used to estimate the diffusion coefficient D , which in turn can be used to check the consistency of the fast-relaxation condition.⁴⁶

IV. THE HELIX-COIL TRANSITION OF DECA-ALANINE

In this section, through an exemplary SMD simulation, we illustrate the PMF calculation method of Sec. III, and demonstrate the Gaussian nature of the resulting work distribution.

We choose as an exemplary system the helix-coil transition of deca-alanine in vacuum. Deca-alanine is an oligopeptide composed of ten alanine residues. In vacuum at room temperature a molecule of deca-alanine folds into a helix. When it is stretched by an external force, the molecule makes a gradual transition to a random coil. For this system the relevant PMF is the free energy profile as a function of the end-to-end distance of the molecule. In an earlier study²⁵ this system was used to assess the accuracy of PMF calculation methods. The PMF was estimated from irreversible

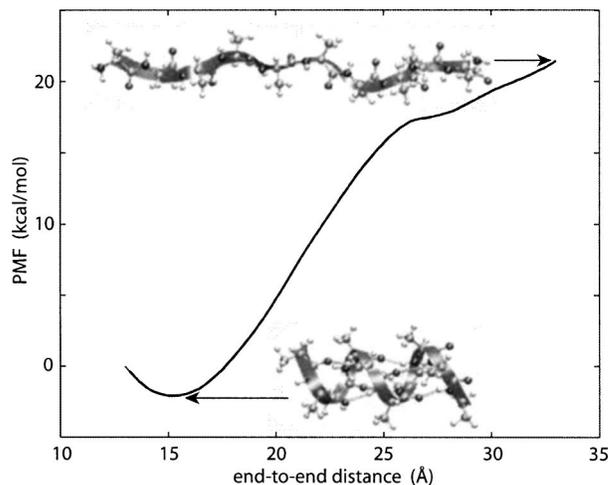


FIG. 3. The PMF of deca-alanine with respect to its end-to-end distance ξ . A typical helical structure at $\xi=15.2$ Å and a typical coil structure at $\xi=33$ Å are shown. The backbones are represented as ribbons. The end-to-end distance is measured between the N atom of the first residue and the capping N atom at the C-terminus. Figure made with VMD (Ref. 47).

(nonequilibrium) stretching simulations through various orders of the cumulant expansion [Eq. (67)] and through the exponential average [Eq. (58)]. The accuracy of the calculated PMFs was assessed compared to the exact PMF obtained from reversible (quasiequilibrium) stretching simulations. Shown in Fig. 3 is the exact PMF obtained from reversible simulations along with two typical configurations of deca-alanine, a helix and a coil.

Here we stretch deca-alanine in an irreversible manner and examine the resulting distribution of work. In the simulation, one end of the molecule (the N atom of the first residue) is fixed at the origin and the other end (the capping N atom at the C-terminus) is constrained to move only along the z axis. The guiding potential $h_\lambda(\mathbf{r}) = (k/2)[\xi(\mathbf{r}) - \lambda]^2$, with the spring constant $k = 500$ pN/Å, is added to control the end-to-end distance ξ . The molecule is stretched by changing the parameter λ from 13 to 33 Å with a constant speed. Two different speeds, 10 and 100 Å/ns, are used. These speeds are, respectively, 100 and 1000 times higher than the reversible speed.²⁵ For the sampling of trajectories, we select initial coordinates from a pool of 10 ns equilibrium simulation (with λ fixed at 13 Å) and initial momenta from the Maxwell-Boltzmann distribution. All simulations were done at constant temperature (300 K) with the temperature controlled by Langevin dynamics. The integration time step of 2 fs was used. The molecular dynamics program NAMD (Ref. 48) was used with the CHARMM22 force field.⁴⁹

The spring constant of 500 pN/Å is large enough to ensure that the end-to-end distance ξ closely follows the constraint center λ .²⁵ For the PMF calculation, we use the leading order, $\Phi(\lambda) = F_\lambda$, in the stiff-spring approximation [Eq. (65)]. The next order is found to be small (<0.5 kcal/mol) compared to the overall scale of the PMF.

A. The work distribution and the PMF calculation

Figures 4 and 5 show analyses of the simulations for $v = 10$ and 100 Å/ns, respectively. For each speed, 10 000 tra-

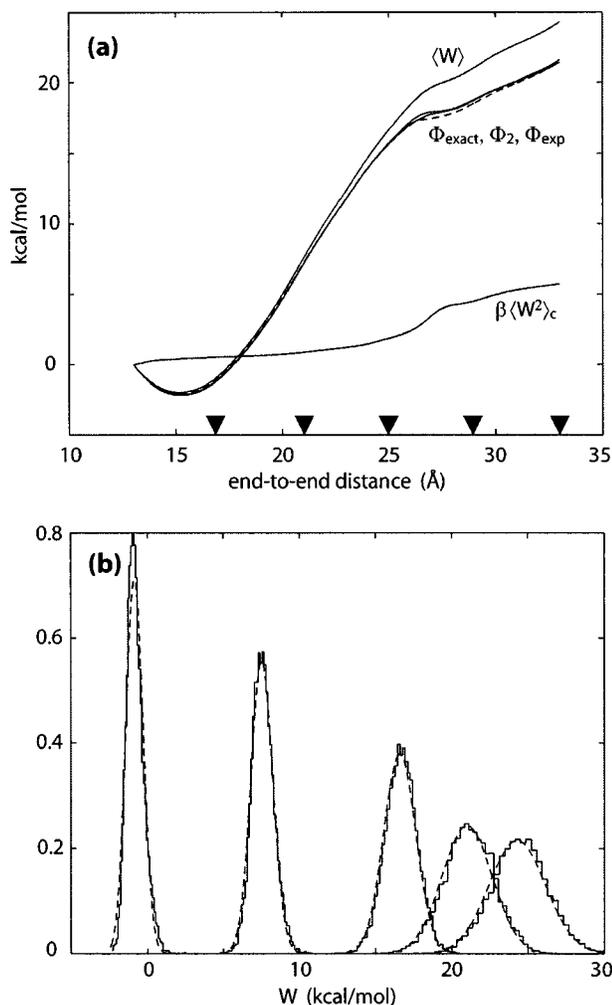


FIG. 4. The PMF calculation and the work distribution from the irreversible stretching simulations with the speed $v = 10 \text{ \AA/ns}$. (a) Plotted against the end-to-end distance are the average work ($\langle W \rangle$), the variance of work ($\langle W^2 \rangle_c$), the PMF estimated with the exponential average (Φ_{exp}), and the PMF estimated with the second order cumulant expansion (Φ_2). The exact PMF Φ_{exact} is shown as a dashed line. (b) Five normalized histograms of work at five different end-to-end distances, $\xi = 17, 21, 25, 29$, and 33 \AA as marked by triangles in (a), are compared with the Gaussian curves (dashed lines) determined from the mean $\langle W \rangle$ and the variance $\langle W^2 \rangle_c$.

jectories were generated. For each trajectory, the work W is calculated as in Eq. (73). The distribution of work indeed seems to be Gaussian throughout the entire course of the process, as can be seen from Figs. 4(b) and 5(b) in which five histograms at five different end-to-end distances [marked by triangles in Figs. 4(a) and 5(a)] are compared with the Gaussian curves determined from the mean $\langle W \rangle$ and the variance $\langle W^2 \rangle_c$. The average work $\langle W \rangle$ includes the irreversible work, which is discounted by Jarzynski's equality. In Figs. 4(a) and 5(a), estimates of the PMF (Φ_{exp} from the exponential average estimator and Φ_2 from the second order cumulant estimator) are compared with the exact PMF Φ_{exact} obtained in Ref. 25. For $v = 10 \text{ \AA/ns}$, both Φ_{exp} and Φ_2 give excellent estimates for the PMF in the entire region; they are almost indistinguishable from Φ_{exact} . For $v = 100 \text{ \AA/ns}$, on the other hand, the estimates are good up to $\xi \approx 25 \text{ \AA}$, but afterward start to diverge from Φ_{exact} ; Φ_{exp} is slightly better for $0 \leq \xi \leq 21 \text{ \AA}$ and Φ_2 is better for the rest of the region.

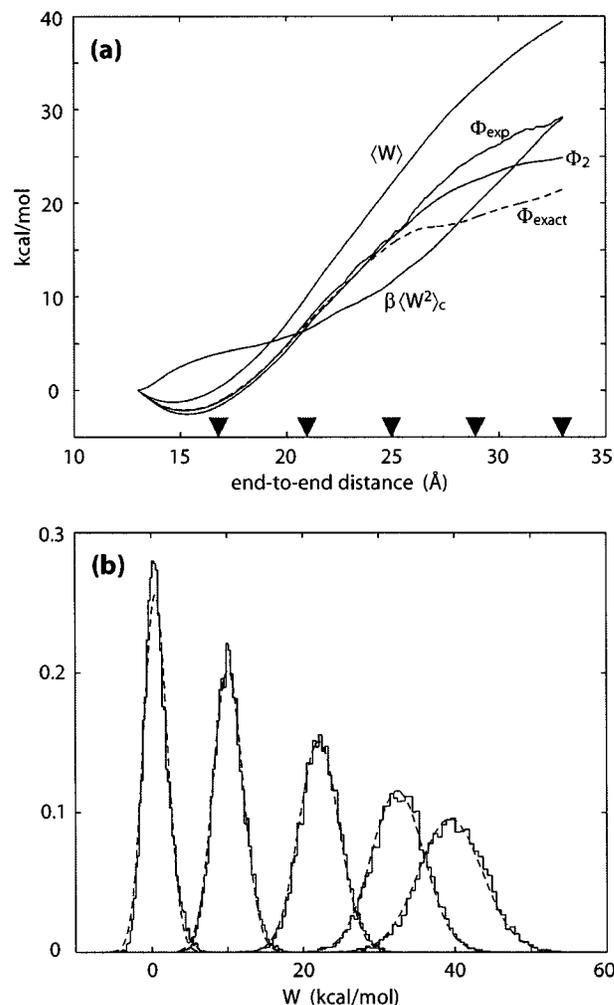


FIG. 5. The PMF calculation and the work distribution from the irreversible stretching simulations with the speed $v = 100 \text{ \AA/ns}$. See the caption of Fig. 4 for details.

Recall that there are two kinds of error involved in the PMF calculation: the truncation error and the sampling error (Sec. III C). As illustrated in Fig. 2 and discussed in Sec. III C, the sampling error generally increases with the work fluctuation $\sqrt{\langle W^2 \rangle_c}$, which in this example grows up to 1.9 kcal/mol ($3.1 k_B T$) for $v = 10 \text{ \AA/ns}$ and 4.2 kcal/mol ($7.0 k_B T$) for $v = 100 \text{ \AA/ns}$. In the ideal case in which the work distribution is perfectly Gaussian and the sampling is perfect, both Φ_{exp} and Φ_2 should be equal to Φ_{exact} . The result for $v = 10 \text{ \AA/ns}$ [Fig. 4(a)] seems to be very close to this ideal situation. However, the result for $v = 100 \text{ \AA/ns}$ [Fig. 5(a)] shows some discrepancy. The discrepancy between Φ_{exp} and Φ_{exact} can be attributed entirely to the sampling error; it would require more trajectories to make Φ_{exp} accurate in the entire region. The discrepancy between Φ_2 and Φ_{exact} is possibly due to both truncation error⁵⁰ and sampling error.

B. Error analysis of the PMF calculation from finite sampling

We needed as many as 10 000 trajectories (for each stretching speed) in order to examine the work distribution.

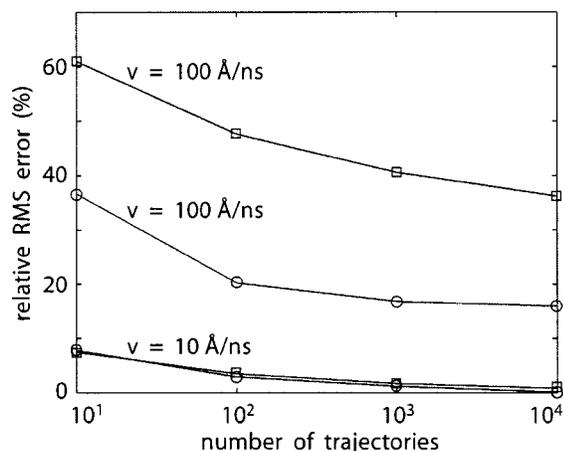


FIG. 6. Error analysis. The relative root-mean-square (RMS) errors for $\Delta\Phi$, the total change in the PMF, estimated with the exponential average estimator (squares) and with the second order cumulant estimator (circles), are shown for four different sampling sizes. The upper two curves correspond to $v = 100 \text{ \AA/ns}$, and the lower two curves to $v = 10 \text{ \AA/ns}$.

It was possible to generate such a large number of trajectories because our system is fairly small (104 atoms). In usual SMD simulations of proteins (typically involving $\sim 10^5$ atoms), however, current computational technology only permits much fewer trajectories. Therefore, it is important to study the accuracy of the PMF calculation method in the case of small sampling sizes.

Using the 10 000 trajectories generated, we examine the accuracy of the estimated PMFs for various sampling sizes. Since the accuracy generally decreases with the stretching distance, we use as a measure of accuracy the relative root-mean-square (RMS) error for the end-point difference in the PMF, $\Delta\Phi \equiv \Phi(33 \text{ \AA}) - \Phi(13 \text{ \AA})$. Four different sampling sizes are considered: 10, 10², 10³, and 10⁴. For each sampling size, all 10 000 trajectories are used. For example, for the sampling size 10, we divide the 10 000 trajectories into 1000 sets of 10 trajectories, estimate $\Delta\Phi$ from each set with the exponential average estimator ($\Delta\Phi_{\text{exp}}$) or with the second order cumulant estimator ($\Delta\Phi_2$), calculate for each set the relative RMS error from the exact value $\Delta\Phi_{\text{exact}} = 21.4 \text{ kcal/mol}$, and take the average of the 1000 errors calculated.

The result is shown in Fig. 6. For $v = 10 \text{ \AA/ns}$, there is only a small difference between the accuracies of $\Delta\Phi_{\text{exp}}$ and $\Delta\Phi_2$; for the sampling size 10 the former is slightly better, and for the other sampling sizes considered the latter is slightly better. For $v = 100 \text{ \AA/ns}$, on the other hand, $\Delta\Phi_2$ gives substantially better estimates; the error of $\Delta\Phi_2$ is only one half of that of $\Delta\Phi_{\text{exp}}$. Overall, the second order cumulant estimator yields the more robust estimate. This finding is somewhat contradictory to the conclusion of Ref. 22.

A common question in computational studies using Jarzynski's equality is how to use optimally a given amount of computing time. Is it advantageous to generate fewer slower trajectories or more faster trajectories? In the present example, we can make three comparisons based on equal amounts of computing time: (i) 10 trajectories of 10 \AA/ns versus 100 trajectories of 100 \AA/ns , (ii) 100 trajectories of 10

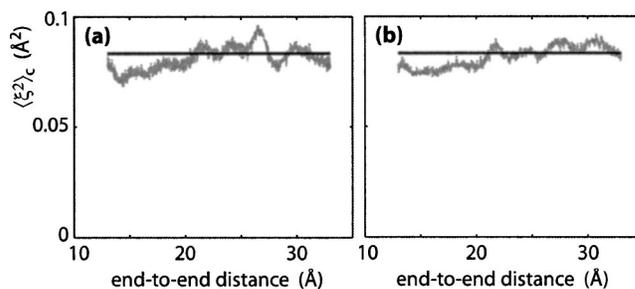


FIG. 7. The variance of the reaction coordinate ξ . (a) $v = 10 \text{ \AA/ns}$. (b) $v = 100 \text{ \AA/ns}$. The straight lines denote the value $1/\beta k$.

\AA/ns versus 1000 trajectories of 100 \AA/ns , and (iii) 1000 trajectories of 10 \AA/ns versus 10 000 trajectories of 100 \AA/ns . As can be seen from Fig. 6, for all these three comparisons fewer slower trajectories win.

C. Time evolution of cumulants and the diffusion coefficient

Under the stiff-spring condition and the assumption of the overdamped Langevin equation, time evolution of the cumulants involving the reaction coordinate ξ and the work W obeys the differential equations given in Eq. (88). When the fast-relaxation condition [Eq. (90)] is satisfied, the solution to these differential equations is given by Eq. (91), or equivalently by Eq. (92). As discussed in Sec. III E, Eqs. (92a) and (92c) are direct consequences of the definition of the work W . Equation (92b) is a statement of the second order formula of Jarzynski's equality, the validity of which was already examined in Secs. IV A and IV B. Now we examine, in the present example of deca-alanine, Eqs. (92d) and (92e) which are consequences of the overdamped Langevin equation.

Figure 7 shows $\langle \xi^2 \rangle_c$, the variance of the reaction coordinate, fluctuating around $1/\beta k$, the value stated in Eq. (92d). Figure 8 shows two curves corresponding to the position-dependent diffusion coefficient estimated with Eq. (92e) from the data for $v = 10$ and 100 \AA/ns , respectively. Although the two curves do not completely coincide, their overall

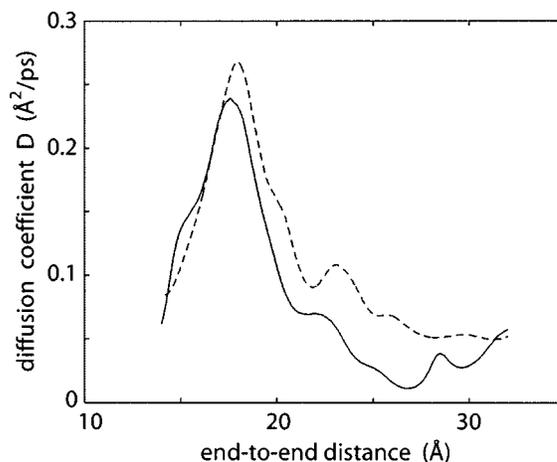


FIG. 8. The position-dependent diffusion coefficient estimated with Eq. (92e). The solid line is from the data for $v = 10 \text{ \AA/ns}$, and the dashed line $v = 100 \text{ \AA/ns}$.

shapes agree. As deca-alanine is stretched from its equilibrium length ($\xi=15.2$ Å), the diffusion coefficient increases, reaches a peak at $\xi\approx 18$ Å, and then decreases again. The estimated diffusion coefficient is in the range of 0.01 Å²/ps $\leq D \leq 0.27$ Å²/ps. Accordingly we find 0.003 Å $\leq v/\beta kD \leq 0.08$ Å for $v=10$ Å/ns and 0.03 Å $\leq v/\beta kD \leq 0.8$ Å for $v=100$ Å/ns. Therefore $v/\beta kD$ is always small compared to the length scale over which the diffusion coefficient changes considerably, validating the fast-relaxation condition [Eq. (90)].

V. CONCLUDING REMARKS

We have discussed theoretical and practical issues concerning the calculation of PMFs from SMD simulations. In particular, we have noticed that, under the stiff-spring condition and the assumption of the overdamped Langevin equation, SMD simulations result in Gaussian work distributions. We have demonstrated the Gaussian nature of work distributions for an exemplary simulation. This result supports the use of the second order cumulant expansion in practical applications of Jarzynski's equality in SMD simulations.

Our method of PMF calculation can be straightforwardly transferred to atomic force microscopy experiments if sufficiently stiff springs are chosen.

ACKNOWLEDGMENTS

We thank Zaida Luthey-Schulten for advice on isobaric–isothermal processes and Chris Jarzynski for general advice and the derivation of Eqs. (42)–(44) in particular, as well as Jin Wang for discussion on the distribution of work. This work was supported by NIH Grants No. PHS-5-P41-RR05969 and No. R01-GM60946 and NSF Grant No. NRAC-MCA93S028.

APPENDIX: VIRIAL EQUATION FOR SYSTEMS UNDER EXTERNAL FORCES

Pressure is the force per unit area exerted by a system on the wall of a container. When the system is in equilibrium, the interaction between the system and the container is balanced with the internal interaction between the constituent particles and the pressure leaves its trace in the internal degrees of freedom. Thus, pressure can be expressed in terms of the internal degrees of freedom, which is the basic idea behind the virial equation. The virial equation is most commonly written as

$$P = \frac{Nk_B T}{V} + \frac{1}{6V} \left\langle \sum_{i \neq j} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} \right\rangle, \quad (\text{A1})$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ is the position of particle i relative to particle j and \mathbf{f}_{ij} is the force on particle i exerted by particle j . A derivation can be found in Ref. 51, Sec. 7.1. The virial equation is particularly useful in computer simulations because it provides a way to calculate pressure without explicitly modeling the interaction between the system and the container. One can also impose certain pressure and simulate the system under that pressure.⁵²

Is this virial equation valid for systems subject to external forces, for example, in SMD simulations in which some

particles are harmonically constrained? Should one include in this case the constraining forces in the calculation of pressure? This question is best answered by tracing the derivation of the virial equation, bearing in mind the more general situation (the presence of external forces).

The virial equation is based on

$$\frac{d}{dt} \left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{p}_i \right\rangle = 0, \quad (\text{A2})$$

where \mathbf{r}_i and \mathbf{p}_i are the position and the momentum of particle i , respectively. This equation is true in equilibrium. In fact, in equilibrium any relevant average is time-independent. The average appearing in Eq. (A2) is just the particular one that leads to the virial equation. An important point is that it must be possible for the system to reach equilibrium. Harmonic constraints applied to some particles will certainly permit equilibrium. However, a uniform external field with a periodic boundary condition will not; in this case the virial equation loses its basis. Therefore we exclude from discussion those cases in which equilibrium is impossible.

Distributing the time derivative, we obtain

$$\left\langle \sum_i \frac{d\mathbf{r}_i}{dt} \cdot \mathbf{p}_i \right\rangle + \left\langle \sum_i \mathbf{r}_i \cdot \frac{d\mathbf{p}_i}{dt} \right\rangle = 0. \quad (\text{A3})$$

We write the force on particle i as the sum of $\mathbf{f}_i^{\text{wall}}$ (the force due to the interaction with the wall) and \mathbf{f}_i (all the other forces including the interparticle forces and external forces such as constraining forces):

$$\frac{d\mathbf{p}_i}{dt} = \mathbf{f}_i + \mathbf{f}_i^{\text{wall}}, \quad \mathbf{f}_i = \sum_{j(\neq i)} \mathbf{f}_{ij} + \mathbf{f}_i^{\text{ext}}. \quad (\text{A4})$$

By using also $d\mathbf{r}_i/dt = \mathbf{p}_i/m_i$, Eq. (A3) becomes

$$\left\langle \sum_i \frac{\mathbf{p}_i^2}{m_i} \right\rangle + \left\langle \sum_i \mathbf{r}_i \cdot \mathbf{f}_i \right\rangle + \left\langle \sum_i \mathbf{r}_i \cdot \mathbf{f}_i^{\text{wall}} \right\rangle = 0. \quad (\text{A5})$$

The third term on the left-hand side can be expressed in terms of pressure as follows. The wall interacts with the system only through the boundary of the system. Thus $\mathbf{f}_i^{\text{wall}}$ is zero unless particle i happens to be on the boundary, and only those particles present on the boundary need to be included in the sum $\sum_i \mathbf{r}_i \cdot \mathbf{f}_i^{\text{wall}}$. We divide the boundary into infinitesimal patches (denoted by α), collect the particles on each patch, and collect all the patches

$$\begin{aligned} \left\langle \sum_i \mathbf{r}_i \cdot \mathbf{f}_i^{\text{wall}} \right\rangle &= \left\langle \sum_{\alpha} \sum_{i \in \alpha} \mathbf{r}_i \cdot \mathbf{f}_i^{\text{wall}} \right\rangle \\ &= \sum_{\alpha} \mathbf{r}(\alpha) \cdot \left\langle \sum_{i \in \alpha} \mathbf{f}_i^{\text{wall}} \right\rangle, \end{aligned} \quad (\text{A6})$$

where $\mathbf{r}(\alpha)$ is the position of patch α . Let us denote the area of patch α by $a(\alpha)$ and the outward normal vector by $\mathbf{n}(\alpha)$. The quantity $\langle \sum_{i \in \alpha} \mathbf{f}_i^{\text{wall}} \rangle$, i.e., the average force exerted on the system by the wall through patch α , is equal to $-Pa(\alpha)\mathbf{n}(\alpha)$. The minus sign means that the force is inward. Substituting this in the preceding equation leads to

$$\left\langle \sum_i \mathbf{r}_i \cdot \mathbf{f}_i^{\text{wall}} \right\rangle = -P \sum_{\alpha} a(\alpha) \mathbf{r}(\alpha) \cdot \mathbf{n}(\alpha) = -P \int_{\partial V} d\mathbf{s} \cdot \mathbf{r}. \quad (\text{A7})$$

In the last step we converted the sum into a surface integral over the boundary of the volume V . From Gauss' theorem follows

$$\int_{\partial V} ds \cdot \mathbf{r} = \int_V dr^3 \nabla \cdot \mathbf{r} = 3V. \quad (\text{A8})$$

Using these results in Eq. (A5), we find

$$P = \frac{1}{3V} \left\langle \sum_i \left(\frac{\mathbf{p}_i^2}{m_i} + \mathbf{r}_i \cdot \mathbf{f}_i \right) \right\rangle. \quad (\text{A9})$$

Recall that \mathbf{f}_i includes external forces as well as the interparticle forces. This form of the virial equation is valid for systems under external forces as long as an equilibrium exists. For pure systems free from external force, we recover the common form [Eq. (A1)] after using Newton's third law ($\mathbf{f}_{ij} = -\mathbf{f}_{ji}$) and replacing $\langle \mathbf{p}_i^2/m_i \rangle$ by its thermal average $3k_B T$.

The virial equation written as in Eq. (A1), in terms of relative positions rather than absolute positions, makes the translational invariance transparent. Naturally, the question arises whether Eq. (A9) is translationally invariant. Upon replacing \mathbf{r}_i by $\mathbf{r}_i + \mathbf{r}_0$, a new term appears

$$P = \frac{1}{3V} \left\langle \sum_i \left(\frac{\mathbf{p}_i^2}{m_i} + \mathbf{r}_i \cdot \mathbf{f}_i \right) \right\rangle + \frac{1}{3V} \mathbf{r}_0 \cdot \left\langle \sum_i \mathbf{f}_i \right\rangle. \quad (\text{A10})$$

In simulations, the quantity $\langle \sum_i \mathbf{f}_i \rangle$ is the average total force on the system because the force from the wall is not explicitly modeled. Since we assume the existence of equilibrium, the average total force must be zero; otherwise there would be a net acceleration. Thus, the additional term indeed vanishes and the virial equation in the form of Eq. (A9) is translationally invariant.

¹A. R. Leach, *Molecular Modelling, Principles and Applications* (Addison-Wesley, Essex, 1996).

²*Computational Biochemistry and Biophysics*, edited by O. M. Becker, A. D. MacKerell, Jr., B. Roux, and M. Watanabe (Marcel Dekker, New York, 2001).

³S. Izrailev, S. Stepaniants, B. Isralewitz, D. Kosztin, H. Lu, F. Molnar, W. Wriggers, and K. Schulten, in *Computational Molecular Dynamics: Challenges, Methods, Ideas*, Vol. 4 in *Lecture Notes in Computational Science and Engineering*, edited by P. Deuffhard, J. Hermans, B. Leimkuhler, A. E. Mark, S. Reich, and R. D. Skeel (Springer-Verlag, Berlin, 1998), pp. 39–65.

⁴B. Isralewitz, M. Gao, and K. Schulten, *Curr. Opin. Struct. Biol.* **11**, 224 (2001).

⁵A. Krammer, H. Lu, B. Isralewitz, K. Schulten, and V. Vogel, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1351 (1999).

⁶M. Gao, M. Wilmanns, and K. Schulten, *Biophys. J.* **83**, 3435 (2002).

⁷M. Gao, D. Craig, V. Vogel, and K. Schulten, *J. Mol. Biol.* **323**, 939 (2002).

⁸S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, and K. Schulten, *Biophys. J.* **72**, 1568 (1997).

⁹M. V. Bayas, K. Schulten, and D. Leckband, *Biophys. J.* **84**, 2223 (2003).

¹⁰C. Jarzynski, *Phys. Rev. Lett.* **78**, 2690 (1997).

¹¹C. Jarzynski, *Phys. Rev. E* **56**, 5018 (1997).

¹²G. E. Crooks, *Phys. Rev. E* **61**, 2361 (2000).

¹³C. Jarzynski, *J. Stat. Phys.* **98**, 77 (2000).

¹⁴T. Hatano and S. Sasa, *Phys. Rev. Lett.* **86**, 3463 (2001).

¹⁵Y. Oono and M. Paniconi, *Prog. Theor. Phys. Suppl.* **130**, 29 (1998).

¹⁶J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, *Science* **296**, 1832 (2002).

¹⁷D. A. Hendrix and C. Jarzynski, *J. Chem. Phys.* **114**, 5974 (2001).

¹⁸G. Hummer and A. Szabo, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3658 (2001).

¹⁹G. Hummer, *J. Chem. Phys.* **114**, 7330 (2001).

²⁰D. M. Zuckerman and T. B. Woolf, *Chem. Phys. Lett.* **351**, 445 (2002).

²¹D. M. Zuckerman and T. B. Woolf, *Phys. Rev. Lett.* **89**, 180602 (2002).

²²J. Gore, F. Ritort, and C. Bustamante, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12564 (2003).

²³M. Ø. Jensen, S. Park, E. Tajkhorshid, and K. Schulten, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6731 (2002).

²⁴R. Amaro, E. Tajkhorshid, and Z. Luthey-Schulten, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7599 (2003).

²⁵S. Park, F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten, *J. Chem. Phys.* **119**, 3559 (2003).

²⁶D. J. Evans and D. J. Searles, *Phys. Rev. E* **50**, 1645 (1994).

²⁷G. Gallavotti and E. G. D. Cohen, *Phys. Rev. Lett.* **74**, 2694 (1995).

²⁸S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks, *J. Chem. Phys.* **103**, 4613 (1995).

²⁹In thermodynamics, the term *isothermal* usually means that the temperature of the system remains constant during the process. Here we use the term to mean that the system is in contact with a bath at constant temperature, though the temperature of the system is not necessarily constant during the process. In fact, the system's temperature is not even defined when we deal with nonequilibrium processes. The same applies to the term *isobaric*.

³⁰Generalization of Jarzynski's equality to the isobaric–isothermal ensemble was discussed in G. E. Crooks, *J. Stat. Phys.* **90**, 1481 (1998).

³¹E. Fermi, *Thermodynamics* (Dover, New York, 1937).

³²C. Jarzynski, *J. Stat. Phys.* **96**, 415 (1999).

³³M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, New York, 1987).

³⁴D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. (Academic, San Diego, 2002).

³⁵C. W. Gardiner, *Handbook of Stochastic Methods*, 2nd ed. (Springer-Verlag, Berlin, 1985).

³⁶There is a third type of method, such as the isokinetic Gaussian thermostat, in which equations of motion are modified by the addition of deterministic terms, but without introducing new degrees of freedom. The nonequilibrium work relation has been established for the isokinetic Gaussian thermostat. See, D. J. Evans, *Mol. Phys.* **101**, 1551 (2003).

³⁷S. Nosé, *J. Chem. Phys.* **81**, 511 (1984).

³⁸W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).

³⁹H. Goldstein, *Classical Mechanics*, 2nd ed. (Addison-Wesley, Reading, 1980).

⁴⁰G. Binnig, C. F. Quate, and C. Gerber, *Phys. Rev. Lett.* **56**, 930 (1986).

⁴¹J. Marcinkiewicz, *Math. Z.* **44**, 612 (1939).

⁴²R. H. Wood, W. C. F. Mühlbauer, and P. T. Thompson, *J. Phys. Chem.* **95**, 6670 (1991).

⁴³J. Hermans, *J. Phys. Chem.* **95**, 9029 (1991).

⁴⁴O. Mazonka and C. Jarzynski, cond-mat/9912121 (1999).

⁴⁵N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).

⁴⁶In fact, any of $\langle \zeta \rangle$, $\langle \Omega \rangle$, $\langle \zeta \Omega \rangle_c$, and $\langle \Omega^2 \rangle_c$ can be used to estimate the diffusion coefficient [see Eq. (91)]. However, data of $\langle \zeta \rangle$ and $\langle \zeta \Omega \rangle_c$ are generally noisier than those of $\langle \Omega \rangle$ and $\langle \Omega^2 \rangle_c$. (Notice that $\langle \Omega \rangle$ and $\langle \Omega^2 \rangle_c$ can be expressed in terms of integrations of $\langle \zeta \rangle$ and $\langle \zeta \Omega \rangle_c$, respectively.) And the use of $\langle \Omega \rangle$ requires the knowledge of the exact PMF. It is the simplest to use $\langle \Omega^2 \rangle_c$, or $\langle W^2 \rangle_c$.

⁴⁷W. Humphrey, A. Dalke, and K. Schulten, *J. Mol. Graphics* **14**, 33 (1996).

⁴⁸L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinzaki, K. Varadarajan, and K. Schulten, *J. Comput. Phys.* **151**, 283 (1999).

⁴⁹A. D. MacKerell, Jr., D. Bashford, M. Bellott *et al.*, *J. Phys. Chem. B* **102**, 3586 (1998).

⁵⁰Recall that the Gaussian nature of the work distribution depends on the assumption of the overdamped Langevin equation (Sec. III D). It is desirable to test this assumption in simulations.

⁵¹M. Plischke and B. Bergersen, *Equilibrium Statistical Mechanics*, 2nd ed. (Word Scientific, Singapore, 1994).

⁵²One way of doing this is the Langevin piston method, which is introduced in Sec. II G.