



Multicanonical molecular dynamics by variable-temperature thermostats and variable-pressure barostats

Cheng Zhang and Michael W. Deem

Citation: *J. Chem. Phys.* **138**, 034103 (2013); doi: 10.1063/1.4773435

View online: <http://dx.doi.org/10.1063/1.4773435>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v138/i3>

Published by the [American Institute of Physics](http://www.aip.org).

Additional information on *J. Chem. Phys.*

Journal Homepage: <http://jcp.aip.org/>

Journal Information: http://jcp.aip.org/about/about_the_journal

Top downloads: http://jcp.aip.org/features/most_downloaded

Information for Authors: <http://jcp.aip.org/authors>

ADVERTISEMENT



Goodfellow
metals • ceramics • polymers • composites
70,000 products
450 different materials
small quantities fast

www.goodfellowusa.com

Multicanonical molecular dynamics by variable-temperature thermostats and variable-pressure barostats

Cheng Zhang^{1,2} and Michael W. Deem^{1,2,3}

¹Applied Physics Program, Rice University, Houston, Texas 77005, USA

²Department of Bioengineering, Rice University, Houston, Texas 77005, USA

³Department of Physics & Astronomy, Rice University, Houston, Texas 77005, USA

(Received 15 November 2012; accepted 12 December 2012; published online 15 January 2013)

Sampling from flat energy or density distributions has proven useful in equilibrating complex systems with large energy barriers. Several thermostats and barostats are presented to sample these flat distributions by molecular dynamics. These methods use a variable temperature or pressure that is updated on the fly in the thermodynamic controller. These methods are illustrated on a Lennard-Jones system and a structure-based model of proteins. © 2013 American Institute of Physics. [<http://dx.doi.org/10.1063/1.4773435>]

I. INTRODUCTION

Molecular dynamics (MD) is a useful tool to study complex molecular systems. A regular MD simply follows Newton's equation, which conserves the total energy E and thus generates configurations in a microcanonical ensemble. We can, however, modify the equation of motion by an artificial thermostat¹⁻⁹ to sample a different distribution, in which the total energy fluctuates, such as the canonical distribution.

For a system with high energy barriers, the limited exploration of the energy landscape in the canonical distribution is insufficient. A multicanonical ensemble,¹⁰⁻¹³ in which the energy is broadly sampled, is more suitable, see Fig. 1(a). The construction of the multicanonical ensemble is nontrivial, for the flat energy distribution requires the inverse of the unknown density of states, $1/\Omega(E)$, as the sampling weight.

We here pursue the construction of multicanonical MD by using a variable temperature equal to $\beta(E) = d \log \Omega / dE$. We describe several thermostats to set this temperature. The method is built on a previous Monte Carlo (MC) method by Yan and de Pablo.¹⁴ Here, the sampling along the total energy is directly incorporated into the variable-temperature thermostat for the MD simulation.

In Sec. II, we describe the method. In Sec. III, we numerically test the method in a few examples. We conclude the article in Sec. IV.

II. METHOD

A. Microcanonical, canonical, and multicanonical ensembles

1. Microcanonical ensemble

In classical statistical mechanics, integration of Newton's equation generates configurations in a microcanonical ensemble under a fixed total energy $E = U(\mathbf{q}) + K(\mathbf{p})$, where $U(\mathbf{q})$ and $K(\mathbf{p}) = \mathbf{p}^2/2m$ are the potential and kinetic energies, respectively. Thus, configurations from an ergodic trajectory

sum to the density of states,

$$\begin{aligned} \Omega(E) &\propto \int \delta[E - U(\mathbf{q}) - K(\mathbf{p})] d\mathbf{q} d\mathbf{p} \\ &= C \int [E - U(\mathbf{q})]^{N_f/2-1} \Theta[E - U(\mathbf{q})] d\mathbf{q}, \end{aligned} \quad (1)$$

where $\Theta(x)$ is the step function, which is 1 if $x > 0$ or 0 otherwise, $C = \sqrt{2m\pi}^{N_f/2} / \Gamma(N_f/2)$, and $\delta(x) = \Theta'(x)$ is the δ -function. We have also integrated out the N_f momentum degrees of freedom,¹⁵

$$\begin{aligned} &\int \delta(E - U - \mathbf{p}^2/2m) d\mathbf{p} \\ &= \frac{\sqrt{2m\pi}^{N_f}}{\Gamma(N_f/2)} (E - U)^{N_f/2-1} \Theta(E - U) \end{aligned}$$

with $\Gamma(x)$ being the gamma function. For a molecular system, N_f is three times the number of particles less the number of conserved quantities, such as the total linear and angular momenta.

The microcanonical temperature, or the derivative of the entropy $S(E) = k_B \log \Omega(E)$ ^{16,17} with k_B being the Boltzmann factor, can be evaluated as a configuration average,^{15,18}

$$\beta(E) \equiv \frac{d \log \Omega(E)}{dE} = \left\langle \frac{N_f/2 - 1}{K(\mathbf{p})} \right\rangle_E, \quad (2)$$

where $\beta(E) = 1/[k_B T(E)]$ is the inverse temperature, $\langle \dots \rangle_E$ denotes an average under a fixed total energy E , and we have used Eq. (1),

$$\begin{aligned} &\frac{1}{\Omega(E)} \frac{d\Omega(E)}{dE} \\ &= \int \left(\frac{N_f/2 - 1}{E - U(\mathbf{q})} \right) \frac{C [E - U(\mathbf{q})]^{N_f/2-1} \Theta[E - U(\mathbf{q})]}{\Omega(E)} d\mathbf{q} \\ &= \left\langle \frac{N_f/2 - 1}{K} \right\rangle_E. \end{aligned}$$

Here, we have assumed $N_f > 2$ so that the singularity at $E = U(\mathbf{q})$ is negligible. Other expressions for $\beta(E)$ also exist,¹⁸ as shown in Appendix A.

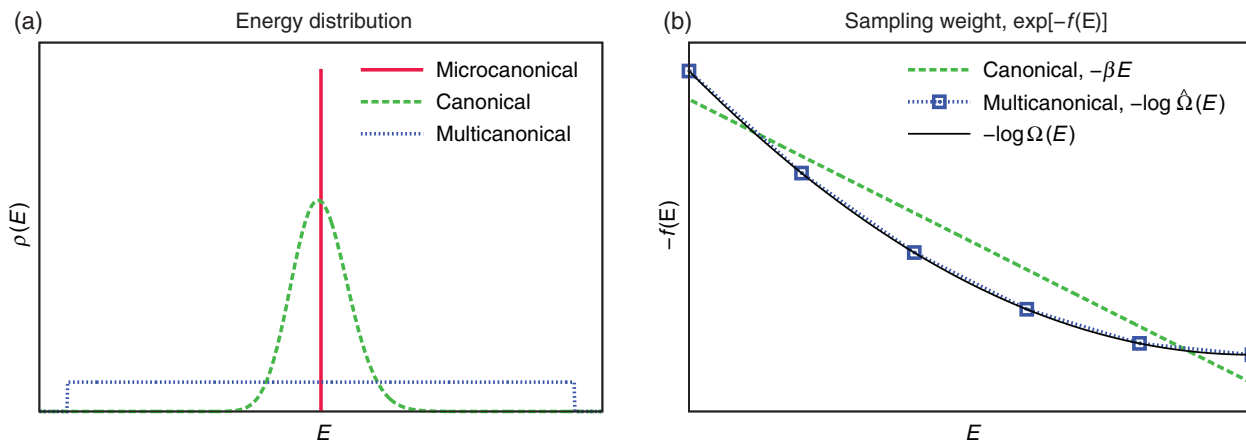


FIG. 1. (a) The probability distributions for the energy in the microcanonical, canonical, and multicanonical ensembles. (b) The ensemble weights in the canonical and microcanonical ensembles.

2. Canonical ensemble

The canonical ensemble is a superposition of microcanonical ensembles at various E , each with a weight $\exp(-\beta E)$. The overall E distribution is given by

$$\begin{aligned} \rho(E) &\propto \int \delta[E - H(\mathbf{q}, \mathbf{p})] d\mathbf{q} d\mathbf{p} \exp(-\beta E) \\ &= \Omega(E) \exp(-\beta E). \end{aligned} \quad (3)$$

To realize a canonical distribution in MD, one needs a thermostat¹⁻⁹ to alter the equation of motion either deterministically³⁻⁸ or stochastically.^{1,2,8,9}

3. Multicanonical ensemble

We now consider a generalized ensemble composed of microcanonical ensembles with an arbitrary weight $\exp[-f(E)]$ instead of the Boltzmann weight $\exp(-\beta E)$. We approximate $f(E)$ by a piecewise linear function on a finely spaced grid E_0, E_1, \dots, E_n , as in Fig. 1(b). Then within each (E_k, E_{k+1}) , $f(E) \approx f(E_k) + f'(E_k)(E - E_k)$ and the ensemble can be treated as locally canonical with the effective temperature being $f'(E_k)$.¹⁰ The constant $f(E_k) - f'(E_k)E_k$ does not distinguish configurations, and is therefore irrelevant. In the limit of infinitesimal intervals, the local effective temperature becomes $f'(E)$.

The multicanonical ensemble is the special case of the generalized ensemble described above, in which $f(E) = \log \Omega(E) + \text{const.}$, such that the E distribution

$$\begin{aligned} \rho_{\text{muca}}(E) &= \int \delta[E - H(\mathbf{q}, \mathbf{p})] d\mathbf{q} d\mathbf{p} \exp[-f(E)] \\ &\propto \Omega(E) \exp[-\log \Omega(E)] = 1 \end{aligned}$$

is flat. Thus, the local temperature $f'(E)$ should be energy dependent, and it is equal to $(\log \Omega)'(E)$, which can be evaluated by Eq. (2).

Thus, molecular dynamics in the multicanonical ensemble can be achieved by a thermostat with an energy-dependent target temperature $f'(E) = (\log \Omega)'(E)$. Below, we describe how this can be done. As in the previous MC method,¹⁴ $\Omega(E)$

is estimated from its derivative by Eq. (2) instead of from the energy histogram.^{10,12,13} Strictly speaking, detailed balance is violated if we constantly update $(\log \Omega)'(E)$ by Eq. (2) in simulation. Earl and Deem, however, showed that this practice, commonly used in the adaptive integration methods^{14,19} and to be used here, satisfies balance in infinite-order Markov chains.²⁰ This updating practice converges in the present application, and so the present method satisfies balance at long times. Note also that the multicanonical ensemble here is formulated for the total energy¹⁴ instead of the potential energy.^{10,12}

B. Algorithm

We now describe an algorithm for a multicanonical MD. The aim is to sample a flat energy distribution in a given range (E_{\min}, E_{\max}) . We split the range into n intervals at $E_0 = E_{\min}, E_1, \dots, E_n = E_{\max}$, such that each (E_k, E_{k+1}) accumulates data for computing the average of $(N_f/2 - 1)/K(\mathbf{p})$ to estimate $\beta(E)$ in Eq. (2),

$$\hat{\beta}(E) = \left\langle \frac{N_f/2 - 1}{K(\mathbf{p})} \right\rangle_{E_k < E < E_{k+1}}. \quad (2')$$

The algorithm is as follows:

1. Integrate Newton's equation for Δt_{MD} .
2. Locate the interval k that contains the current total energy E (if $E_{\min} < E < E_{\max}$) and deposit $(N_f/2 - 1)/K(\mathbf{p})$ into the accumulator there.
3. Use Eq. (2') to estimate $\hat{\beta}(E)$; but if the number of data points is fewer than 100, or if E lies outside of (E_{\min}, E_{\max}) , use a default value $\hat{\beta}(E) = \beta_d$, which should correspond to an energy $E \in (E_{\min}, E_{\max})$.
4. Use $\hat{\beta}(E)$ in one of the variable-temperature thermostats listed below.
5. Repeat Step 1 until the simulation ends.

1. Remarks

1. We show in Appendix A that the use of the interval-averaged formula (2') to approximate $\beta(E)$ produces no systematic error: the estimated values of the density of

states $\log \hat{\Omega}(E_k) = \int^{E_k} \hat{\beta}(E) dE$ at the boundaries E_k asymptotically approach $\log \Omega(E_k)$ for any interval size $E_{k+1} - E_k$.

- While the average in Eq. (2') is invariant with respect to the interval size, the energy range, the number of intervals, and the number of data points will affect the rate of convergence. The number of intervals should be large enough so that the $\log \hat{\Omega}(E)$ is sufficiently close to $\log \Omega(E)$, i.e., the dotted line segments in Fig. 1(b) closely approximate the solid curve. Furthermore, the number of data points per interval should be sufficiently large to provide a reliable estimate of the average.
- In Step 3, a constant $\hat{\beta}(E) = \beta_d$ is assumed for E outside of (E_{\min}, E_{\max}) for simplicity. It, however, causes a large discontinuity there, and a better boundary condition may improve this. The number 100 is also arbitrary.
- After simulation, we can recover the density of potential-energy state $g(U) = \int \delta[U - U(\mathbf{q})] d\mathbf{q}$ from the observed distribution $h(U)$ of the potential energy U by

$$g(U) = h(U) / \left[C \int_U^{+\infty} (E - U)^{N_f/2 - 1} / \hat{\Omega}(E) dE \right], \quad (4)$$

see Appendix B for a proof. The denominator of Eq. (4) can be evaluated exactly by the incomplete gamma function,²¹ as $\log \hat{\Omega}(E)$ is piecewise linear. One can then compute the canonical partition function $Z(\beta) = \int g(U) \exp(-\beta U) dU$, and its derivatives. Equation (4) also applies to an $\hat{\Omega}(E)$ that does not produce a flat E distribution, see examples in Appendix A.

- The algorithm produces a flat total-energy E distribution, but not a flat potential-energy U distribution. To roughly achieve the latter, we note $\rho(E) = \rho_{\langle U \rangle_E}(\langle U \rangle_E) d\langle U \rangle_E / dE$, where

$$\begin{aligned} d\langle U \rangle_E / dE &= \left(\frac{1}{2} N_f - 1 \right) (\langle K \rangle_E \langle 1/K \rangle_E - 1) \\ &= - \left(\frac{1}{2} N_f - 1 \right) \langle \delta K \delta(1/K) \rangle_E \\ &= \left(\frac{1}{2} N_f - 1 \right) \langle \delta U \delta(1/K) \rangle_E. \end{aligned}$$

This factor would enter the variable-temperature thermostats described below either through the target density of states $\hat{\Omega}_{\text{mod}}(E) = \hat{\Omega}(E) / (d\langle U \rangle_E / dE)$, or, equivalently, through the target inverse temperature $\hat{\beta}_{\text{mod}}(E) = \hat{\beta}(E) - d \log(d\langle U \rangle_E / dE) / dE$. This modification would lead to a flat $\rho_{\langle U \rangle_E}(\langle U \rangle_E)$. In the thermodynamic limit, the overall potential-energy distribution $h(U)$ would also be flat, as $h(U) \approx \rho_{\langle U \rangle_E}(U)$. For finite systems, $h(U)$ may be flatter with $\hat{\Omega}_{\text{mod}}(E)$ and $\hat{\beta}_{\text{mod}}(E)$ than with $\hat{\Omega}(E)$ and $\hat{\beta}(E)$.

C. Variable-temperature thermostats

The thermostats for a canonical ensemble need to be modified as follows for a multicanonical MD. We list a few examples below. The dimension is denoted by d , which is usually 3.

1. Nosé-Hoover chain thermostat

The Nosé-Hoover chain equations^{5,8} for a canonical ensemble are

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{p}/m, \quad \dot{\mathbf{p}} = \mathbf{F} - \zeta_1 \mathbf{p}, \quad \dot{\zeta}_1 = (\mathbf{p}^2/m - N_f/\beta) / Q_1 - \zeta_2 \zeta_1, \\ \dot{\zeta}_k &= (Q_{k-1} \zeta_{k-1}^2 - 1/\beta) / Q_k - \zeta_{k+1} \zeta_k, \quad (\text{for } k=2, \dots, M-1), \\ \dot{\zeta}_M &= (Q_{M-1} \zeta_{M-1}^2 - 1/\beta) / Q_M, \end{aligned} \quad (5)$$

where ζ_k are the thermostat variables and Q_k are their masses. The corresponding Liouville's equation

$$\partial \rho / \partial t + \partial(\dot{\mathbf{q}} \rho) / \partial \mathbf{q} + \partial(\dot{\mathbf{p}} \rho) / \partial \mathbf{p} + \sum_{k=1}^M \partial(\dot{\zeta}_k \rho) / \partial \zeta_k = 0 \quad (6)$$

has a stationary ($\partial \rho / \partial t = 0$) solution

$$\rho(\mathbf{q}, \mathbf{p}, \{\zeta_k\}) \propto \exp \left[-\beta U(\mathbf{q}) - \beta \mathbf{p}^2 / 2m - \beta \sum_{k=1}^M Q_k \zeta_k^2 / 2 \right],$$

which is indeed canonical.⁵

For a multicanonical ensemble, we change the equation for $\dot{\zeta}_1$ in Eq. (5) to

$$\dot{\zeta}_1 = \{ [\hat{\beta}(E) / \beta_0] \mathbf{p}^2 / m - N_f / \beta_0 \} / Q_1 - \zeta_2 \zeta_1, \quad (5')$$

and we replace the β in the equations for $\dot{\zeta}_k$, $k \geq 2$, by β_0 , such that the new stationary solution of Eq. (6) is

$$\begin{aligned} \rho(\mathbf{q}, \mathbf{p}, \{\xi_k\}) &\propto \exp \left\{ -\log \hat{\Omega}[U(\mathbf{q}) + \mathbf{p}^2 / 2m] \right. \\ &\quad \left. - \beta_0 \sum_{k=1}^M Q_k \xi_k^2 / 2 \right\}, \end{aligned} \quad (7)$$

where β_0 is a reference inverse temperature, against which the set point of the system's inverse temperature is scaled in the term in braces in Eq. (5'). That is, with the use of $\hat{\beta}(E)$ in Eq. (5'), the stationary distribution is the desired Eq. (7), no matter the value of β_0 . One expects that β_0 values near the average of $\hat{\beta}(E)$ will be efficient for sampling.

We recover the original Nosé-Hoover thermostat by setting $\zeta_2 = 0$, and removing the equations for ζ_k with $k \geq 2$ and for ζ_M .

2. Velocity-rescaling thermostat

The velocity-rescaling thermostat⁹ achieves a canonical distribution by re-sampling the distribution of the total energy E under fixed \mathbf{q} coordinates,

$$\rho(E) dE \propto [E - U(\mathbf{q})]^{N_f/2 - 1} \exp(-\beta E) dE,$$

cf. Eqs. (1) and (3). The change $E \rightarrow E'$, hence $K \rightarrow K'$, is realized by scaling the velocity or momentum: $\mathbf{p} \rightarrow \sqrt{K'/K} \mathbf{p}$, for $U(\mathbf{q})$ is fixed. The kinetic energy is randomly updated and follows the Langevin equation⁹

$$dK/dt = \frac{1}{2} N_f / \beta - K + \sqrt{2K/\beta} \xi, \quad (8)$$

where ξ is a Gaussian white noise that satisfies $\langle \xi(t) \xi(t') \rangle = \delta(t - t')$.

For a multicanonical ensemble, we modify Eq. (8) to

$$dK/dt = \frac{1}{2} N_f / \beta_0 - [\hat{\beta}(E) / \beta_0] K + \sqrt{2K/\beta_0} \xi, \quad (8')$$

such that the corresponding Fokker-Planck equation

$$\partial\rho/\partial t + \partial\left(\left\{\frac{1}{2}N_f - [\hat{\beta}(E)/\beta_0]K\right\}\rho\right)/\partial K - \partial^2(K\rho/\beta_0)/\partial K^2 = 0$$

has the desired stationary solution $\rho(K) \propto K^{N_f/2-1}/\hat{\Omega}(U+K)$. Here, β_0 is a reference inverse temperature.

Note that Eq. (8) or (8') can be integrated with a different time step from the MD one.

3. Monte-Carlo velocity-rescaling thermostat

We can also replace Eq. (8') by an explicit MC move in the E or K space.¹⁴ We first propose a change of $\log K$ by $\delta \in (-\Delta, \Delta)$, then accept the new $K' = Ke^\delta$ by

$$A(K \rightarrow K') = \min \left\{ 1, \frac{\hat{\Omega}(K+U)}{\hat{\Omega}(K'+U)} \left(\frac{K'}{K}\right)^{N_f/2} \right\}.$$

If K' is accepted, it is also realized as a velocity rescaling: $\mathbf{p} \rightarrow \sqrt{K'/K}\mathbf{p}$.

4. Langevin dynamics

Langevin dynamics^{1,17} modifies Newton's equation of motion by introducing a velocity damping term and a thermal noise,

$$\dot{\mathbf{q}} = \mathbf{p}/m, \dot{\mathbf{p}} = \mathbf{F} - \zeta\mathbf{p} + \sqrt{2m\zeta/\beta}\xi, \quad (9)$$

where ζ is a damping constant and ξ is a Gaussian white noise satisfying $\langle \xi_i(t)\xi_j(t') \rangle = \delta(t-t')\delta_{ij}$.

For a multicanonical ensemble, we change Eq. (9) to

$$\dot{\mathbf{q}} = \mathbf{p}/m, \dot{\mathbf{p}} = \mathbf{F} - [\hat{\beta}(E)/\beta_0]\zeta\mathbf{p} + \sqrt{2m\zeta/\beta_0}\xi, \quad (9')$$

such that the corresponding Fokker-Planck equation

$$\partial\rho/\partial t + \partial[(\mathbf{p}/m)\rho]/\partial\mathbf{q} + \partial\{[\mathbf{F} - (\hat{\beta}(E)/\beta_0)\zeta\mathbf{p}]\rho\}/\partial\mathbf{p} - \partial^2[(m\zeta/\beta_0)\rho]/\partial\mathbf{p}^2 = 0$$

has the desired stationary ($\partial\rho/\partial t = 0$) solution $\rho(\mathbf{q}, \mathbf{p}) \propto 1/\hat{\Omega}[U(\mathbf{q}) + \mathbf{p}^2/2m]$. Here, β_0 is a reference inverse temperature.

5. Andersen thermostat

In the Andersen thermostat,^{2,8} we replace the velocity or momentum $\mathbf{p}_i^{\text{old}}$ of a random particle i by a Gaussian random vector drawn from the Maxwell distribution,

$$p_{\text{Maxwell}}(\beta; \mathbf{p}_i^{\text{new}}) = (\beta/2\pi m)^{d/2} \exp[-\beta(\mathbf{p}_i^{\text{new}})^2/2m]. \quad (10)$$

For a multicanonical ensemble, we treat $\mathbf{p}_i^{\text{new}}$ from Eq. (10), with β replaced by $\hat{\beta}(E)$, as a MC trial, and accept it by the Metropolis probability

$$A(\mathbf{p}_i^{\text{old}} \rightarrow \mathbf{p}_i^{\text{new}}) = \min \left\{ 1, \frac{\hat{\Omega}(E) p_{\text{Maxwell}}(\hat{\beta}(E'); \mathbf{p}_i^{\text{old}})}{\hat{\Omega}(E') p_{\text{Maxwell}}(\hat{\beta}(E); \mathbf{p}_i^{\text{new}})} \right\}, \quad (11)$$

where $E' = E + (\mathbf{p}_i^{\text{new}})^2/2m - (\mathbf{p}_i^{\text{old}})^2/2m$ and the two Maxwell factors $p_{\text{Maxwell}}(\dots)$, defined in Eq. (10), offset the *a priori* sampling bias. For the canonical ensemble, $\hat{\beta}(E) = \hat{\beta}(E') = \beta$, $\hat{\Omega}(E)/\hat{\Omega}(E') = \exp[\beta(\mathbf{p}_i^{\text{old}})^2/2m - \beta(\mathbf{p}_i^{\text{new}})^2/2m]$, and Eq. (11) yields unity so that the trial is always accepted.

D. Volume distribution

We can extend the above sampling strategy to explore a broad volume distribution, which is useful in studying a liquid-gas phase transition.^{22,23} A flat volume distribution can be inefficient, however, as drastic free energy changes occur only at a high-density, or small-volume, regime. We therefore sample the distribution $\rho(V) \propto 1/V^\alpha$. To sample a flat density distribution, $\rho_n(n) = \text{const.}$, with $n = N/V$, we use $\alpha = 2$, because

$$\rho(V)dV = \rho_n(n)|dn| = \rho_n(n)NdV/V^2 \propto (1/V^2)dV.$$

To sample the distribution $\rho(V) \propto V^{-\alpha}$, we replace the constant target pressure in an isothermal-isobaric simulation^{2,5,7,8,24-26} by the variable

$$\hat{p}(V) = \frac{\langle p_{\text{int.}}(\mathbf{q}, \mathbf{p})V^\alpha \rangle_{V_k < V < V_{k+1}}}{\langle V^\alpha \rangle_{V_k < V < V_{k+1}}}, \quad (12)$$

where $p_{\text{int.}}(\mathbf{q}, \mathbf{p}) = (\mathbf{p}^2/m + \mathbf{F} \cdot \mathbf{q})/(d \cdot V) - d\phi_{\text{tail}}(V)/dV$, where $\phi_{\text{tail}}(V)$ is the energy tail correction, such that the potential energy $U(\mathbf{q}, V) = U_{\text{trunc.}}(\mathbf{q}) + \phi_{\text{tail}}(V)$.^{7,8,24-26} The weight V^α in Eq. (12) ensures unbiased free energy $\hat{F}(V_k) = -\int_{V_{\text{min}}}^{V_k} \hat{p}(V)dV$ for a finite interval width $V_{k+1} - V_k$, see Appendix A.

1. Nosé-Hoover chain barostat

To sample $\rho(V) \propto V^{-\alpha}$, we modify the Nosé-Hoover chain^{5,7,8} equation for an isothermal-isobaric ensemble to

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{p}/m + \eta\mathbf{q}, \dot{\mathbf{p}} = \mathbf{F} - (\zeta + \eta)\mathbf{p}, \dot{V} = d\eta V, \\ \dot{\eta} &= \{[p_{\text{int.}}(\mathbf{q}, \mathbf{p}) - \hat{p}(V)]V + (1 - \alpha)/\beta\}d/W - \zeta_1\eta, \\ \dot{\zeta}_1 &= [\mathbf{p}^2/m + W\eta^2 - (N_f + 1)/\beta]/Q_1 - \zeta_2\zeta_1, \\ \dot{\zeta}_k &= (Q_{k-1}\zeta_{k-1}^2 - 1/\beta)/Q_k - \zeta_{k+1}\zeta_k, \text{ (for } k=2, \dots, M-1), \\ \dot{\zeta}_M &= (Q_{M-1}\zeta_{M-1}^2 - 1/\beta)/Q_M, \end{aligned} \quad (13)$$

where η and ζ_k are the variables for the barostat and thermostat, respectively, with W and Q_k being their respective masses. The corresponding Liouville's equation

$$\begin{aligned} \partial\rho/\partial t + \partial(\dot{\mathbf{q}}\rho)/\partial\mathbf{q} + \partial(\dot{\mathbf{p}}\rho)/\partial\mathbf{p} + \partial(\dot{\eta}\rho)/\partial\eta \\ + \sum_{k=1}^M \partial(\dot{\zeta}_k\rho)/\partial\zeta_k + \partial(\dot{V}\rho)/\partial V = 0 \end{aligned}$$

has the desired stationary solution

$$\begin{aligned} \rho(\mathbf{q}, \mathbf{p}, \eta, \{\zeta_k\}, V) \propto \exp \left\{ -\beta[U(\mathbf{q}, V) + \mathbf{p}^2/2m + W\eta^2/2 \right. \\ \left. + \sum_{k=1}^M Q_k\zeta_k^2/2 - \hat{F}(V)] \right\} V^{-\alpha}, \end{aligned}$$

where $\hat{F}(V) = -\int^V \hat{p}(V')dV'$ is the estimated Helmholtz free energy. The term $1-\alpha$ in the equation for $\dot{\eta}$ accounts for the ensemble weight $V^{-\alpha}$. We recover an isothermal-isobaric sampling by setting $\hat{p}(V)$ to a constant and $\alpha \rightarrow 0$. The $1-\alpha \rightarrow 1$ corrects the small error in the original Nosé-Hoover barostat.^{5,7,8}

We recover the Nosé-Hoover case by setting $\zeta_2 = 0$, and removing the equations for $\dot{\zeta}_k$ with $k \geq 2$ and for $\dot{\zeta}_M$.

2. Langevin barostat

Alternatively, the pressure can be controlled by a Langevin equation²⁶ separate from the thermostat. The Langevin barostat is a stochastic version of Eq. (13),

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{p}/m + \eta\mathbf{q}, \quad \dot{\mathbf{p}} = \mathbf{F} - \eta\mathbf{p}, \quad \dot{V} = d\eta V, \\ \dot{\eta} &= \{[p_{\text{int}}(\mathbf{q}, \mathbf{p}) - \hat{p}(V)]V + (1-\alpha)/\beta\}d/W - \zeta\eta \\ &\quad + \sqrt{2\zeta/(\beta W)}\xi, \end{aligned} \quad (14)$$

where ξ is a Gaussian white noise satisfying $\langle \xi(t)\xi(t') \rangle = \delta(t-t')$, and ζ and W are two constants. The Fokker-Planck equation

$$\begin{aligned} \partial\rho/\partial t + \partial(\dot{\mathbf{q}}\rho)/\partial\mathbf{q} + \partial(\dot{\mathbf{p}}\rho)/\partial\mathbf{p} + \partial(\dot{V}\rho)/\partial V \\ + \partial[(\dot{\eta})_{\text{det.}}\rho]/\partial\eta - \partial^2[(\zeta/\beta W)\rho]/\partial\eta^2 = 0, \end{aligned}$$

where $(\dot{\eta})_{\text{det.}} = \{[p_{\text{int}}(\mathbf{q}, \mathbf{p}) - \hat{p}(V)]V + (1-\alpha)/\beta\}d/W - \zeta\eta$, admits the desired stationary distribution,

$$\begin{aligned} \rho(\mathbf{q}, \mathbf{p}, \eta, V) \propto \exp\{-\beta[U(\mathbf{q}, V) \\ + \mathbf{p}^2/2m + W\eta^2/2 - \hat{F}(V)]\}V^{-\alpha}. \end{aligned}$$

3. Explicit Monte-Carlo sampling

The above Langevin equation can be replaced by a MC move.^{8,22} we first propose a change of $\log V^8$ by $\delta \in (-\Delta, \Delta)$ and then accept the new $V' = Ve^{\delta}$ by

$$A(V \rightarrow V') = \min \left\{ 1, \frac{\exp\{-\beta[U(\sqrt{V'/V}\mathbf{q}, V') + (\sqrt{V'/V}\mathbf{p})^2/2m - \hat{F}(V')]\}V'^{-\alpha}}{\exp\{-\beta[U(\mathbf{q}, V) + \mathbf{p}^2/2m - \hat{F}(V)]\}V^{-\alpha}} \right\}. \quad (15)$$

If the move is accepted, we change the coordinates as $\mathbf{q} \rightarrow \sqrt{V'/V}\mathbf{q}^8$ and momenta as $\mathbf{p} \rightarrow \sqrt{V'/V}\mathbf{p}$. The phase-space volume element $d\mathbf{q}d\mathbf{p}$ is conserved.

III. EXAMPLES

A. A Flat energy distribution

We tested the method in the $N = 108$ particle Lennard-Jones system.⁸ The potential was cutoff at $r_c = 2.5 \text{ \AA}$ and shifted to zero. The density was $n = N/V = 0.7$. Each trajectory was integrated for 5×10^8 steps and the time step was $\Delta t_{\text{MD}} = 0.001$ unless specified otherwise. The energy E range was $(-4N, 2N)$ and the interval was $0.02N$. The default inverse temperature β_d for E outside of the range or insufficient data and reference inverse temperature β_0 were both 1.0.

The energy distributions $\rho(E)$ from simulations using several variable-temperature thermostats are shown in Figs. 2(a)–2(e). The flatness of the energy distributions $\rho(E)$ helped detect if the variable-temperature thermostat was correct. The modifications in Sec. II C were necessary to produce flat distributions. That is, if the β was simply replaced by $\hat{\beta}(E)$ in Eqs. (5), (8), (9), and (10), flat energy distributions were not achieved, as shown in Figs. 2(a), 2(c), 2(e), and 2(f), respectively.

The thermostat equations are stiff if the mass Q_1 in the Nosé-Hoover chain is not extensive. An efficient Q_1 should be proportional to the number N_f of degrees of freedom,^{4,6,25} which can be seen from the scales of the equation for $\dot{\zeta}_1$ in Eq. (5'). From the left side, we have $\rho(\zeta_1)$

$\propto \exp(-\beta_0 Q_1 \zeta_1^2/2)$ by Eq. (7). Thus, $\zeta_1 \propto 1/\sqrt{Q_1}$ and $\dot{\zeta}_1 \sim \zeta_1/\Delta t_{\text{MD}} \sim 1/(\Delta t_{\text{MD}}\sqrt{Q_1})$. From the right side of Eq. (5'), $[\hat{\beta}(E)/\beta_0]\mathbf{p}^2/m - N_f/\beta_0 \sim \sqrt{N_f}$, since it is proportional to the standard deviation of \mathbf{p}^2/m . Thus, the right side is $\sim \sqrt{N_f}/Q_1$. Equating the two results, we find $\sqrt{Q_1}/\Delta t_{\text{MD}} \sim \sqrt{N_f}$. The other Q_k , $k \geq 2$, should be of order unity.^{6,25} With $Q_1 = 10$, which was much smaller than N_f , visible deviations from the flat distribution can be seen, but a smaller MD time step Δt_{MD} recovered a flat distribution, see Fig. 2(f), as it must by Eq. (7).

The approximate¹⁵ temperature formula $\tilde{\beta}(E) = N_f/\langle 2K \rangle_E$ ^{14,27} did not produce a flat energy distribution, but rather a $\rho(E) \propto 1/\langle K \rangle_E$, as shown in Fig. 2(d). This is because $\tilde{\beta}(E)$ corresponds to a modified density of states weighted by $2K$, as discussed in Appendix A.

B. Structure-based protein model

As an application of multicanonical MD, we studied the protein villin headpiece, Protein DataBank code: 1VII,^{28,29} by a simplified structure-based model.^{30–32} Details of the model are described in Appendix C. Each trajectory was run for 2×10^9 steps with $\Delta t_{\text{MD}} = 0.002$. We used the MC velocity-rescaling variable-temperature thermostat with $\Delta = 0.05$. We show below that the thermodynamics of the model depends critically on the cutoff distance r_c that defines contacts; the interaction is attractive between contact atoms but repulsive otherwise.

As shown in Fig. 3(a), the microcanonical temperature functions $\beta(E)$ under several different r_c were

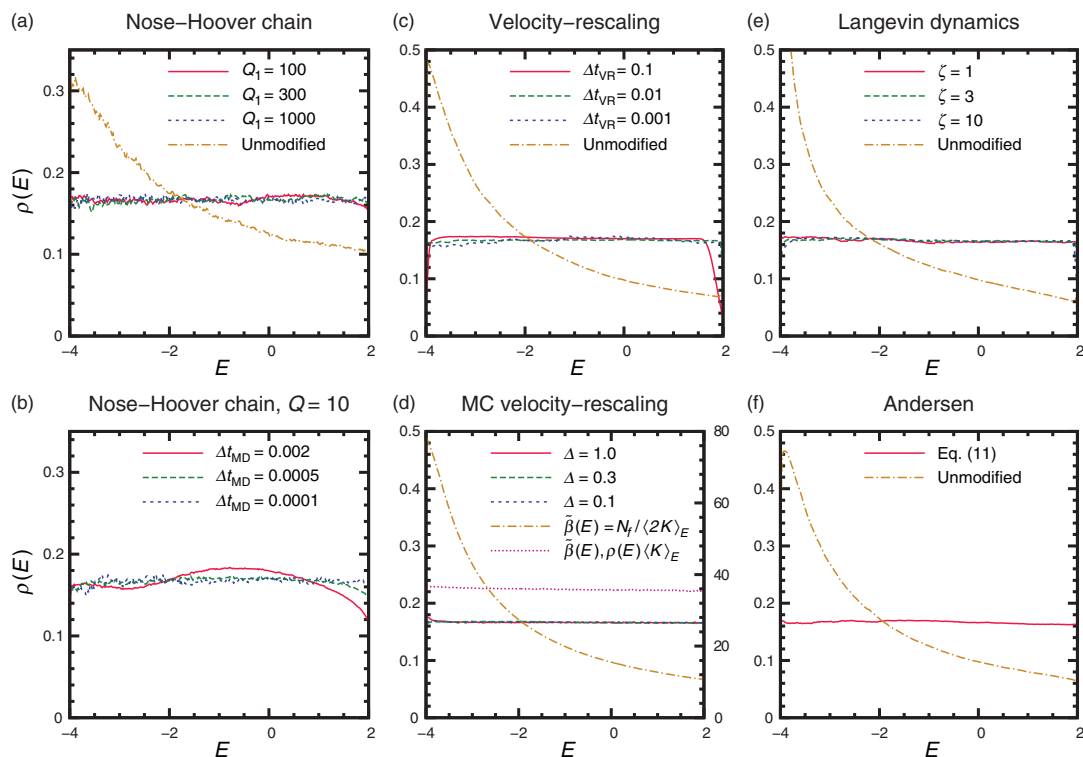


FIG. 2. Multicanonical molecular dynamics on a Lennard-Jones system by the variable-temperature thermostats discussed in Sec. II C. (a)–(f) The probability distributions for the energy in the system for different methods. These variable-temperature thermostats produce flat energy distributions. Simple replacement of β with $\hat{\beta}(E)$ in the canonical thermostat equations, labeled “Unmodified,” does not produce flat energy distributions. The parameters are (a) $Q_1 = 300$, (c) $\Delta_{VR} = 0.01$, and (e) $\zeta = 3$. For the Nosé-Hoover chain variable-temperature thermostat in (a) and (b), $M = 5$ and $Q_k = 1$ for $k \geq 2$. Note in (b) the importance of a small time step if the thermostat mass Q_1 is too small. In panel (d), the alternative expression $\tilde{\beta}(E) = N_f / \langle 2K \rangle_E$ with $\Delta = 0.3$ produced a non-flat energy distribution, characterized in Appendix A.

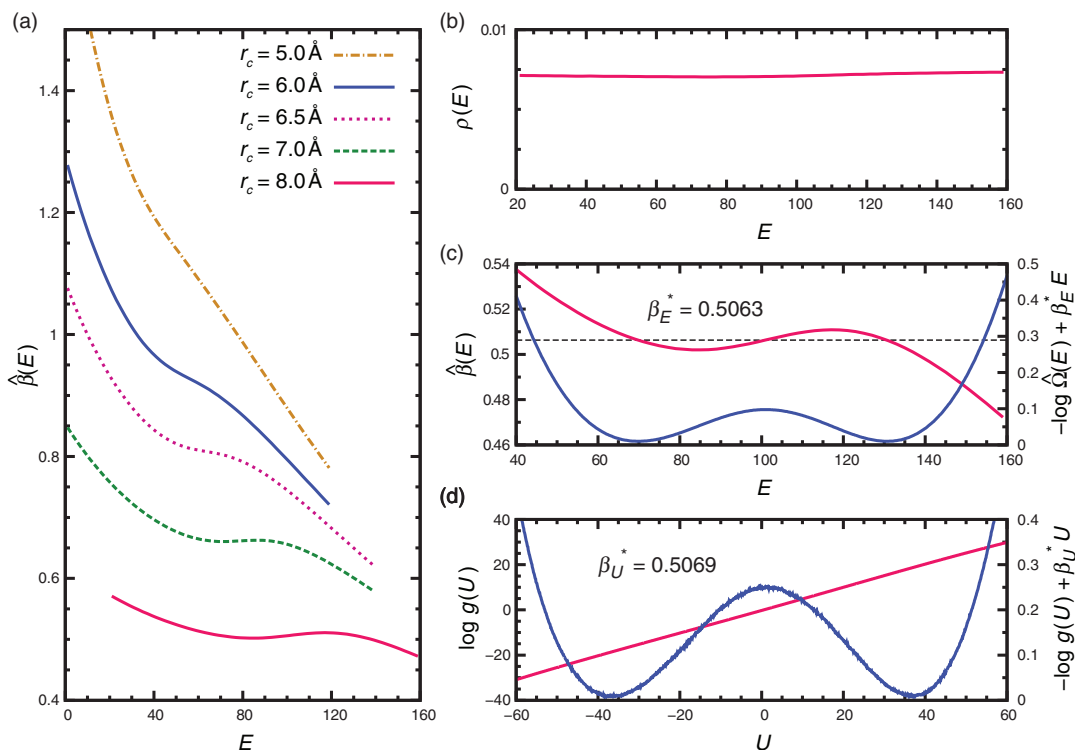


FIG. 3. Simulations on the villin headpiece by a structure-based model. (a) Estimated inverse temperature, $\hat{\beta}(E)$, as a function of energy for different cutoffs, r_c . The default inverse temperature β_d and reference inverse temperature β_0 were the same, and they were set to 1.111 for $r_c = 5 \text{ \AA}$, 0.909 for $r_c = 6 \text{ \AA}$, 0.833 for $r_c = 6.5 \text{ \AA}$, 0.667 for $r_c = 7 \text{ \AA}$, and 0.5 for $r_c = 8 \text{ \AA}$. (b) The probability distribution for the energy in the system. (c) The values of $\hat{\beta}(E)$ (solid, red), β_E^* (dashed), and $-\log \hat{\Omega}(E) + \beta_E^* E$ (dotted, blue). (d) The values of $\log g(U)$ (solid, red) and $-\log g(U) + \beta_U^* U$ (dotted, blue). Panels (b)–(d) were obtained from the simulation with $r_c = 8 \text{ \AA}$.

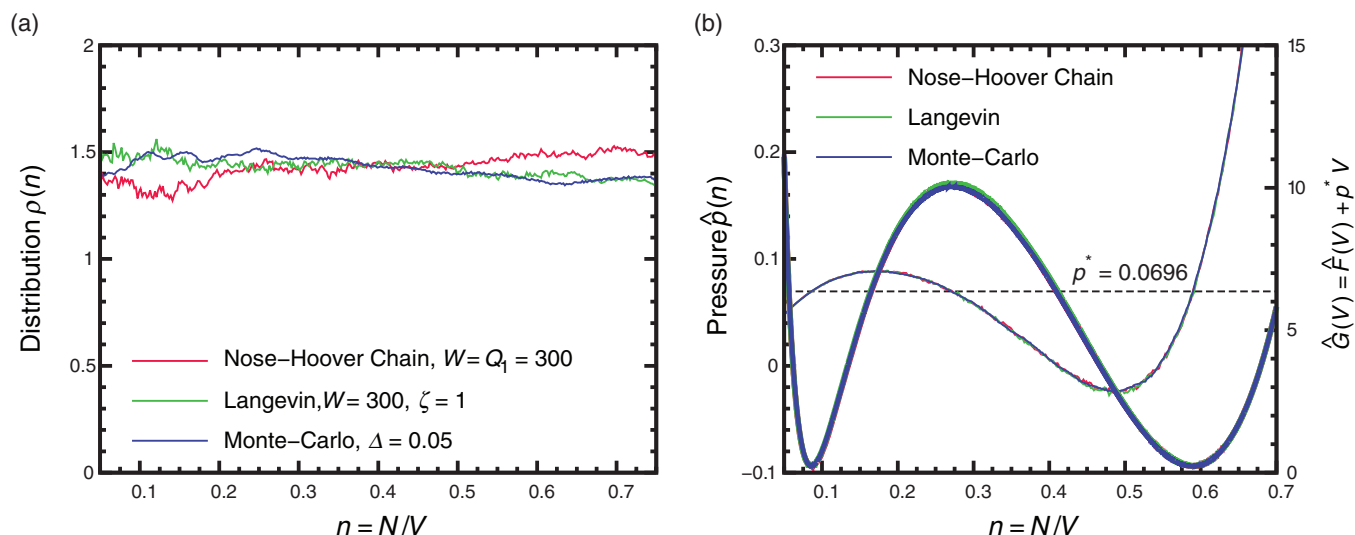


FIG. 4. Molecular dynamics with a flat density distribution on a Lennard-Jones system. (a) Flat density distributions for three different variable-pressure barostats. (b) Values of $\hat{p}(V)$ (thin lines), p_V^* (dashed), and $\hat{G}(V) = \hat{F}(V) + p_V^* V$ (thick lines) for the three barostats. For the Langevin and MC variable-pressure barostats, velocity-rescaling was used as the independent thermostat with $\Delta t_{VR} = 0.01$.

similar to pressure-volume isotherms around a liquid-gas transition in a simple fluid.^{17,33} The back-bending S -loop³⁴ that signifies a first-order phase transition occurred only when $r_c \geq 7$ Å. The sensitivity to r_c may explain that both the presence and absence of S -loop were previously observed in similar models.^{13,35} A strong first-order transition in the energy space may be absent with the common definitions of contacts,^{31,32,36} whose equivalent r_c were about 4 to 5 Å.

As shown in Fig. 3(b), the total-energy distribution was flat. For larger models with higher energy barriers, the modification described in Remark 5 of Sec. II B, which will approximately produce a flat potential-energy histogram, may be more suitable.

For $r_c > 6.5$ Å, the energy distributions $\rho(E)$ around the transition temperature β_E^* were bimodal. Here, β_E^* was defined as the temperature at which the two modes share the same height. See Fig. 3(c) for the case of $r_c = 8$ Å. The free energy profile $-\log \hat{\Omega}(E) + \beta_E^* E = \int^E [\beta_E^* - \hat{\beta}(E')] dE'$ indeed had two basins.

We computed the density of potential-energy state $g(U)$ by the reweighting formula (4). The transition temperature β_U^* determined from $g(U)$ was similar to β_E^* , see Fig. 3(d). The free energy profile $-\log g(U) + \beta_U^* U$ along U showed a higher barrier than that along $-\log \hat{\Omega}(E) + \beta_E^* E$.

C. Flat density distribution

We tested the extension in Sec. II D for the flat density distribution also on the 108 particle Lennard-Jones system. The pair potential was cutoff at half-box length to facilitate volume change moves.⁸ The density range was (0.05, 0.75) with a spacing 0.002. The default pressure p_d for a density outside of the range or insufficient data was 0.1. $T = 1.2$. Each trajectory was run for 4×10^8 steps with $\Delta t_{MD} = 0.001$.

The density distribution, pressure-density curves, and free energy profile are shown in Fig. 4 for the Nosé-Hoover chain equation (13), Langevin equation (14), and MC equa-

tion (15) barostats. The density distributions were flat. The obtained transition pressures were 0.0696, 0.0695, and 0.0697 for the Nosé-Hoover chain, Langevin, and MC barostats, respectively. The free energy profiles obtained at the transition pressures also agreed with one another, see Fig. 4(b).

IV. DISCUSSION AND CONCLUSION

A multicanonical, or flat-energy-distribution, molecular dynamics can be conveniently realized by a variable-temperature thermostat. Similarly, a flat-density-distribution MD can be realized by a variable-pressure barostat. The method is easy to implement with minimal modifications from a regular constant temperature or volume simulation. We have demonstrated the method on a Lennard-Jones system and a structure-based model of the villin headpiece.

The method computes the density of states $\Omega(E)$ ¹⁴ instead of the density of potential-energy states $g(U)$.^{10,12} The reweighting formula (4) is used to recover $g(U)$. In this way, thermal averages in the canonical ensemble are computed after simulation.

The method is somewhat more robust, in terms of achieving a flat energy distribution, with a smaller energy interval $\Delta E = E_{k+1} - E_k$, although it is formally correct for any ΔE , as shown in Appendix A. A smaller interval also yields a finer $\hat{\Omega}(E_k)$, but may slow down the on the fly construction of the ensemble weight. The integral-identity technique,^{29,37} which allows borrowing data from neighboring intervals, and the adaptive averaging technique from the continuous tempering²⁹ may help improve convergence.

The multicanonical ensemble is commonly used in Monte Carlo and molecular dynamics simulations to overcome energy barriers. We here have described a number of variable-temperature thermostats and variable-pressure barostats to achieve flat energy or density distributions.³⁸ We expect these methods may prove useful in simulating systems with complex energy landscapes.

ACKNOWLEDGMENTS

This work was supported by the Shared University Grid at Rice funded by the National Science Foundation (NSF) under Grant No. EIA-0216467 and by the U.S. Department of Energy Grant No. DE-FG02-03ER15456.

APPENDIX A: UNBIASED DENSITY OF STATES AT INTERVAL BOUNDARIES

We show here that the values of the density of states $\hat{\Omega}(E_k) = \int^{E_k} \hat{\beta}(E) dE$ at the boundary points E_k are unbiased, although $\hat{\beta}(E)$ is an approximate average of $\beta(E)$ over finite intervals (E_k, E_{k+1}) . This is because²⁹

$$\begin{aligned} \frac{\Omega(E_{k+1})}{\hat{\Omega}(E_{k+1})} - \frac{\Omega(E_k)}{\hat{\Omega}(E_k)} &= \int_{E_k}^{E_{k+1}} \frac{d}{dE} \left[\frac{\Omega(E)}{\hat{\Omega}(E)} \right] dE \\ &= \int_{E_k}^{E_{k+1}} [(\log \Omega)'(E) - (\log \hat{\Omega})'(E)] \frac{\Omega(E)}{\hat{\Omega}(E)} dE \\ &= h_k \left\langle \frac{N_f/2 - 1}{K(\mathbf{p})} - \hat{\beta}(E) \right\rangle_{E_k < E < E_{k+1}}, \end{aligned}$$

where we have used $(\log \Omega)'(E) = \langle (N_f/2 - 1)/K(\mathbf{p}) \rangle_E$ and $(\log \hat{\Omega})'(E) = \hat{\beta}(E)$. The ratio $\Omega(E)/\hat{\Omega}(E) = \int \delta[E - H(\mathbf{q}, \mathbf{p})] d\mathbf{q} d\mathbf{p} / \hat{\Omega}(E)$ is the distribution density of the overall E distribution, and we have defined $h_k \equiv \int_{E_k}^{E_{k+1}} \Omega(E)/\hat{\Omega}(E) dE$. Since the right hand side is zero by Eq. (2'), the left side vanishes. So $\hat{\Omega}(E_{k+1})/\hat{\Omega}(E_k) = \Omega(E_{k+1})/\Omega(E_k)$, and by recursion, $\hat{\Omega}(E_k)/\hat{\Omega}(E_0) = \Omega(E_k)/\Omega(E_0)$. Thus, $\hat{\Omega}(E_k)$ differs from $\Omega(E_k)$ only by a multiple.

The alternative formulas for $\hat{\beta}(E)$: $\hat{\beta}^{(\mathbf{p})}(E) = N_f / (2K(\mathbf{p}))_{E_k < E < E_{k+1}}$ and $\hat{\beta}^{(\mathbf{F})}(E) = \langle \nabla^2 U \rangle_{E_k < E < E_{k+1}} / \langle \mathbf{F} \cdot \mathbf{F} \rangle_{E_k < E < E_{k+1}}$ in Refs. 14 and 27, have a similar property, although they do not produce flat E distributions, even with infinitely small intervals. To see this, consider an arbitrary vector field $\mathbf{v} = \mathbf{v}(\mathbf{p}, \mathbf{q})$. We define

$$\Omega_D(E) \equiv \int (\mathbf{v} \cdot \nabla H) \delta[E - H(\mathbf{q}, \mathbf{p})] d\mathbf{q} d\mathbf{p} = \Omega(E) \langle \mathbf{v} \cdot \nabla H \rangle_E,$$

$$\Omega_N(E) \equiv \int (\nabla \cdot \mathbf{v}) \delta[E - H(\mathbf{q}, \mathbf{p})] d\mathbf{q} d\mathbf{p} = \Omega(E) \langle \nabla \cdot \mathbf{v} \rangle_E.$$

Now integration by parts yields $d\Omega_D(E)/dE = \Omega_N(E)$ ¹⁸ and

$$\begin{aligned} \frac{\Omega_D(E_{k+1})}{\hat{\Omega}(E_{k+1})} - \frac{\Omega_D(E_k)}{\hat{\Omega}(E_k)} &= \int_{E_k}^{E_{k+1}} \left[\frac{\Omega_N(E)}{\Omega(E)} - (\log \hat{\Omega})'(E) \frac{\Omega_D(E)}{\Omega(E)} \right] \frac{\Omega(E)}{\hat{\Omega}(E)} dE \\ &= h_k [\langle \nabla \cdot \mathbf{v} \rangle_{E_k < E < E_{k+1}} - \hat{\beta}(E) \langle \mathbf{v} \cdot \nabla H \rangle_{E_k < E < E_{k+1}}]. \end{aligned}$$

Thus, if we set $\hat{\beta}(E) = \langle \nabla \cdot \mathbf{v} \rangle_{E_k < E < E_{k+1}} / \langle \mathbf{v} \cdot \nabla H \rangle_{E_k < E < E_{k+1}}$, $\hat{\Omega}(E_k)$ will asymptotically approach the modified density of states $\Omega_D(E_k)$ instead of $\Omega(E_k)$. Accordingly, the converged energy distribution $\rho(E) \propto \Omega(E)/\Omega_D(E) = 1/\langle \mathbf{v} \cdot \nabla H \rangle_E$ is not necessarily flat, and the resulting $\hat{\Omega}(E)$ should be divided by $\langle \mathbf{v} \cdot \nabla H \rangle_E$ in order to recover $\Omega(E)$.

The two alternative mean force formulas mentioned above correspond to $\mathbf{v}^{(\mathbf{p})} \equiv \mathbf{p}$ with $\mathbf{v}^{(\mathbf{p})} \cdot \nabla H = \mathbf{p}^2/m$ and $\mathbf{v}^{(\mathbf{F})} \equiv \mathbf{F}$ with $\mathbf{v}^{(\mathbf{F})} \cdot \nabla H = -\mathbf{F} \cdot \mathbf{F}$, respectively. Thus,

the overall energy distributions $\rho^{(\mathbf{p})}(E) \propto 1/\langle \mathbf{p}^2/m \rangle_E$ and $\rho^{(\mathbf{F})}(E) \propto 1/\langle \mathbf{F} \cdot \mathbf{F} \rangle_E$ are not flat.

We can force a flat distribution by using a ‘‘normalized’’ vector field. For an arbitrary \mathbf{v} , we construct $\mathbf{v}' = \mathbf{v}/(\mathbf{v} \cdot \nabla H)$, then¹⁸

$$\hat{\beta}(E) = \langle \nabla \cdot \mathbf{v}' \rangle_{E_k < E < E_{k+1}} = \langle \nabla \cdot [\mathbf{v}/(\mathbf{v} \cdot \nabla H)] \rangle_{E_k < E < E_{k+1}}$$

yields the unbiased $\hat{\Omega}(E_k)$ and a flat distribution, since $\langle \mathbf{v}' \cdot \nabla H \rangle_E = 1$ and $\Omega_D(E) = \Omega(E)$. The above two examples for $\mathbf{v}^{(\mathbf{p})} = \mathbf{p}$ and $\mathbf{v}^{(\mathbf{F})} = \mathbf{F}$, after the normalization, become Eq. (2') and $\hat{\beta}(E) = \langle \nabla \cdot [\mathbf{F}/(\mathbf{F} \cdot \nabla U)] \rangle_{E_k < E < E_{k+1}}$,¹⁸ respectively.

We can show Eq. (12) in a similar way. We define the canonical partition function at a fixed volume V as

$$Q(V) = \exp[-\beta F(V)] = \int \exp[-\beta H_V(\mathbf{q}, \mathbf{p})] d\mathbf{q} d\mathbf{p},$$

and its estimate

$$\log \hat{Q}(V) = -\beta \hat{F}(V) = \beta \int^V \hat{p}(V') dV'.$$

Then

$$\begin{aligned} \frac{Q(V_{k+1})}{\hat{Q}(V_{k+1})} - \frac{Q(V_k)}{\hat{Q}(V_k)} &= \int_{V_k}^{V_{k+1}} \frac{d}{dV} \left[\frac{Q(V)}{\hat{Q}(V)} \right] dV \\ &= -\beta \int_{V_k}^{V_{k+1}} [F'(V) - \hat{F}'(V)] V^\alpha \frac{Q(V)}{\hat{Q}(V)} \frac{dV}{V^\alpha} \\ &= -\beta \int_{V_k}^{V_{k+1}} [F'(V) - \hat{F}'(V)] V^\alpha w(V) dV \\ &= \beta h_k \langle [p(V) - \hat{p}(V)] V^\alpha \rangle_{V_k < V < V_{k+1}}, \end{aligned}$$

where $w(V) \equiv \int \exp[-\beta H_V(\mathbf{q}, \mathbf{p})] d\mathbf{q} d\mathbf{p} / [\hat{Q}(V) V^\alpha] = Q(V) / [\hat{Q}(V) V^\alpha]$ is the overall volume distribution, and we have defined $h_k \equiv \int_{V_k}^{V_{k+1}} w(V) dV$. Thus, in order for $\hat{Q}(V_k)$ to converge to $Q(V_k)$, we need $\hat{p} = \langle p(V) V^\alpha \rangle_{V_k < V < V_{k+1}} / \langle V^\alpha \rangle_{V_k < V < V_{k+1}}$, which becomes Eq. (12) by $p(V) = \langle p_{\text{int}}(\mathbf{q}, \mathbf{p}) \rangle_V$.

APPENDIX B: REWEIGHTING FOR THE DENSITY OF POTENTIAL-ENERGY STATES

To prove Eq. (4), we first write down the distribution density of the potential energy $h(U)$ in the multicanonical ensemble,

$$\begin{aligned} h(U) &= \int_{-\infty}^{+\infty} \left(\int \delta[E - U(\mathbf{q}) - K(\mathbf{p})] \right. \\ &\quad \left. \times \delta[U - U(\mathbf{q})] d\mathbf{q} d\mathbf{p} \right) [1/\hat{\Omega}(E)] dE. \end{aligned}$$

So

$$h(U) = \int_{-\infty}^{+\infty} \left(C \int [E - U(\mathbf{q})]^{N_f/2-1} \Theta[E - U(\mathbf{q})] \right)$$

$$\begin{aligned}
& \times \delta[U - U(\mathbf{q})]d\mathbf{q}) [1/\hat{\Omega}(E)]dE \\
& = \int_{-\infty}^{+\infty} \left(\int \delta[U - U(\mathbf{q})]d\mathbf{q} \right) \\
& \quad \times C(E - U)^{N_f/2-1} \Theta(E - U) / \hat{\Omega}(E) dE \\
& = g(U) \int_U^{+\infty} C(E - U)^{N_f/2-1} / \hat{\Omega}(E) dE,
\end{aligned}$$

where we have integrated out momenta in the first step, replaced $U(\mathbf{q})$ by U in the second, and used $g(U) = \int \delta[U - U(\mathbf{q})]d\mathbf{q}$ in the last. Dividing both sides by the integral yields Eq. (4).

APPENDIX C: STRUCTURE-BASED PROTEIN MODEL

The protein model was described in Refs. 31 and 32; we include it here for completeness. The model includes alpha-carbon atoms only, and builds a potential energy based on the geometry of a given protein structure,

$$\begin{aligned}
U & = \sum_{\text{bonds}} k_b (b - b_0)^2 + \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 \\
& + \sum_{\substack{\text{dihedrals} \\ n=1,3}} k_\varphi^{(n)} \{1 - \cos[n(\varphi - \varphi_0)]\} \\
& + \sum_{\substack{\text{contacts} \\ i < j - 3}} \varepsilon \left[5 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 6 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{10} \right] \\
& + \sum_{\substack{\text{notcontacts} \\ i < j - 3}} \varepsilon \left(\frac{C}{r_{ij}} \right)^{12},
\end{aligned}$$

where b and b_0 are the bond lengths of successive atoms along the chain in the given configuration and the reference values, respectively; similarly, θ and θ_0 are the angles of three successive atoms, and φ and φ_0 are the dihedrals of four successive atoms. For non-bonded interactions, we distinguish pairs of contact and non-contact atoms; in the former case, we apply the fourth term with σ_{ij} being the distance between i and j in the reference structure, while in the latter case, we apply the last term. We used the following parameters: $k_b = 100$, $k_\theta = 20$, $k_\varphi^{(1)} = 1$, $k_\varphi^{(3)} = 0.5$, and $C = 4.0 \text{ \AA}$.

We used a simpler criterion than the original one^{31,36} to define contact atoms: in the reference structure, if any two non-hydrogen atoms from two different residues have a distance $r_{ij} < r_c$, the corresponding alpha-carbon atoms are contacts. Note that Kouza *et al.*³² used the distances between alpha-carbon, instead of non-hydrogen, atoms to define contacts. The contacts produced by our criterion appeared to be more similar to those yielded by the CSU server.³⁶

¹H. Mori, *Prog. Theor. Phys.* **34**, 399 (1965); R. Zwanzig, *J. Stat. Phys.* **9**, 215 (1973).

- ²H. C. Andersen, *J. Chem. Phys.* **72**, 2384 (1980).
³S. Nosé, *Mol. Phys.* **52**, 255 (1984); C. Braga and K. P. Traversa, *J. Chem. Phys.* **123**, 134101 (2005).
⁴S. Nosé, *J. Chem. Phys.* **81**, 511 (1984).
⁵W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).
⁶G. J. Martyna, M. L. Klein, and M. Tuckerman, *J. Chem. Phys.* **97**, 2635 (1992).
⁷G. J. Martyna, D. J. Tobias, and M. L. Klein, *J. Chem. Phys.* **101**, 4177 (1994).
⁸D. Frenkel and B. Smit, *Understanding Molecular Simulation From Algorithms to Applications*, 2nd ed. (Academic, 2002).
⁹G. Bussi, D. Donadio, and M. Parrinello, *J. Chem. Phys.* **126**, 014101 (2007).
¹⁰B. A. Berg and T. Neuhaus, *Phys. Rev. Lett.* **68**, 9 (1992); J. Lee, *ibid.* **71**, 211 (1993).
¹¹P. M. C. de Oliveira, T. J. P. Penna, and H. J. Herrmann, *Eur. Phys. J. B* **1**, 205 (1998); J. S. Wang, *ibid.* **8**, 287 (1999); J.-S. Wang and R. H. Swendsen, *J. Stat. Phys.* **106**, 245 (2002); N. Nakajima, H. Nakamura, and A. Kidera, *J. Phys. Chem. B* **101**, 817 (1997).
¹²F. G. Wang and D. P. Landau, *Phys. Rev. Lett.* **86**, 2050 (2001); J. Kim, J. E. Straub, and T. Keyes, *ibid.* **97**, 050601 (2006).
¹³J. Kim, J. E. Straub, and T. Keyes, *J. Chem. Phys.* **126**, 135101 (2007).
¹⁴Q. Yan and J. J. de Pablo, *Phys. Rev. Lett.* **90**, 035701 (2003).
¹⁵V. Martin-Mayor, *Phys. Rev. Lett.* **98**, 137207 (2007).
¹⁶L. D. Landau and E. M. Lifshits, *Statistical Physics*, 3rd ed. (Pergamon, Oxford, 1980); S.-K. Ma, *Statistical Mechanics* (World Scientific, Philadelphia, 1985).
¹⁷D. A. McQuarrie, *Statistical Mechanics* (Harper & Row, New York, 1976).
¹⁸H. H. Rugh, *Phys. Rev. Lett.* **78**, 772 (1997); B. D. Butler, G. Ayton, O. G. Jepps, and D. J. Evans, *J. Chem. Phys.* **109**, 6519 (1998); O. G. Jepps, G. Ayton, and D. J. Evans, *Phys. Rev. E* **62**, 4757 (2000).
¹⁹M. Fasnacht, R. H. Swendsen, and J. M. Rosenberg, *Phys. Rev. E* **69**, 056704 (2004).
²⁰D. J. Earl and M. W. Deem, *J. Phys. Chem. B* **109**, 6701 (2005).
²¹W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1992).
²²Q. Yan, T. S. Jain, and J. J. de Pablo, *Phys. Rev. Lett.* **92**, 235701 (2004).
²³C. Zhang and J. Ma, *Phys. Rev. E* **76**, 036708 (2007).
²⁴G. Bussi, T. Zykova-Timan, and M. Parrinello, *J. Chem. Phys.* **130**, 074101 (2009).
²⁵G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein, *Mol. Phys.* **87**, 1117 (1996).
²⁶S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks, *J. Chem. Phys.* **103**, 4613 (1995).
²⁷M. Chopra and J. J. de Pablo, *J. Chem. Phys.* **124**, 114102 (2006); N. Rathore, T. A. Knotts, and J. J. de Pablo, *Biophys. J.* **85**, 3963 (2003).
²⁸Y. Duan and P. A. Kollman, *Science* **282**, 740 (1998); P. L. Freddolino and K. Schulten, *Biophys. J.* **97**, 2338 (2009); K. Lindorff-Larsen, S. Piana, R. O. Dror, and D. E. Shaw, *Science* **334**, 517 (2011).
²⁹C. Zhang and J. Ma, *J. Chem. Phys.* **132**, 244101 (2010).
³⁰N. Go, *Annu. Rev. Biophys. Bioeng.* **12**, 183 (1983); M. Kouza and U. H. E. Hansmann, *J. Chem. Phys.* **134**, 044124 (2011).
³¹C. Clementi, H. Nymeyer, and J. N. Onuchic, *J. Mol. Biol.* **298**, 937 (2000); H. Lammert, A. Schug, and J. N. Onuchic, *Proteins: Struct., Funct., Bioinf.* **77**, 881 (2009).
³²M. Kouza, C.-F. Chang, S. Hayryan, T.-h. Yu, M. S. Li, T.-h. Huang, and C.-K. Hu, *Biophys. J.* **89**, 3353 (2005).
³³J. K. Johnson, J. A. Zollweg, and K. E. Gubbins, *Mol. Phys.* **78**, 591 (1993).
³⁴J. Kim and J. E. Straub, *J. Chem. Phys.* **133**, 154101 (2010).
³⁵M.-H. Hao and H. A. Scheraga, *J. Phys. Chem.* **98**, 4940 (1994); J. Kim, J. E. Straub, and T. Keyes, *Phys. Rev. E* **76**, 011913 (2007).
³⁶V. Sobolev, A. Sorokine, J. Prilusky, E. E. Abola, and M. Edelman, *Bioinformatics* **15**, 327 (1999).
³⁷A. B. Adib and C. Jarzynski, *J. Chem. Phys.* **122**, 014114 (2005).
³⁸See <http://www.mwdeem.rice.edu/tbstat> for the code for thermostats and barostats.