Asymmetric bias in free-energy perturbation measurements using two Hamiltonian-based models

Di Wu and David A. Kofke
Department of Chemical and Biological Engineering, University at Buffalo, The State University of New York, Buffalo, New York 14260-4200, USA

(Received 9 September 2004; published 2 December 2004)

We present two model systems that are suitable for the study of bias in free-energy perturbation (FEP) calculations which are performed in molecular simulations. The models exhibit the asymmetry that is sometimes seen in these calculations, in which the magnitude of the bias is greater when the calculation is performed in one direction ($A \rightarrow B$, sampling system A and perturbing into system B) versus the other ($B \rightarrow A$). Both models are formulated as a system of N independent particles, each sampling a space in the presence of a one-body field that is different for the A and B systems. In one model the field is a harmonic potential. The other model is discrete, such that each particle can be at one of two points (or states) of different energy. The neglected-tail bias model is applied to each system to estimate the average bias as a function of the amount of FEP sampling, and numerical calculations are performed to show that the bias model is effective. We show that the bias is significantly smaller when sampling is performed on the system having a broader work distribution (we designate this direction "insertion") compared to the bias for FEP calculations that sample the system with a narrower distribution ("deletion").

DOI: 10.1103/PhysRevE.70.066702 PACS number(s): 05.10.−a, 65.40.Gr, 05.70.Ln, 02.50.−r

I. INTRODUCTION

Free-energy perturbation (FEP) is a molecular simulation method for evaluating free-energy differences [1]. It entails the simulation of one system to measure an ensemble average involving energy changes upon instantaneous perturbation into a second system. The systems (labeled A and B) may differ in many ways, such as the thermodynamic state or the definition of their Hamiltonians. The importance of FEP methods derives from the central role that free energy plays in characterizing the equilibrium of a system [2]. Measurements of free-energy differences can be used to assess the relative stability of the two systems, thereby furthering understanding of their behavior and enabling their intelligent manipulation.

Many examples can be given. Evaluation of the solubility of a gas or solid in a liquid is fundamental to many applications, and its conduct by molecular simulation involves the evaluation of the chemical potential of one solute molecule at infinite dilution in the liquid. This requires a free-energy calculation for the process in which the molecule is added to (or removed from) the solvent. Particle removal is well known to perform poorly in this application, but particle insertion is not without its complications either [3]. The best way to proceed for a given system is not always obvious, and considerable extra effort can be needed to ensure a good result. The process could be improved with better understanding of the nature of the calculation. As another example, assessment of the relative stability of crystal polymorphs [4] is an important problem from many standpoints: scientifically, technologically, and economically. The question of relative stability can in principle be assessed by a comparison of the free energies of the competing forms. In practice the true free energy is not considered when molecular modeling is applied to this question. Instead, crude approximations such as neglect of entropy are applied. The inability to describe polymorphic behavior reliably is, at least in part, due to the difficulty of evaluating the free energy accurately for these model systems.

In these systems and many others of interest, the problems involved in calculating the free energy have common origins. Sampling must be performed in a way that both systems involved in the free-energy difference are sampled well. Failure to sample well does not always give obviously wrong results and, consequently, it is as important to identify the presence of a sampling problem as it is to develop ways to overcome it. Progress can be made by working with simple, analytically tractable models that exhibit the essential features that give rise to difficulties found in real free-energy calculations and measurements. In studying these models we can develop metrics that signal the presence of a problem with the calculation and that can be used to identify or formulate methods that are effective in a given situation. Presently there are no such models available for this purpose, and this work aims to address this deficiency.

The working equation for a FEP calculation is [5]

$$\exp(-\beta\Delta F) = \langle \exp(-\beta W) \rangle_A, \tag{1}$$

where $\Delta F = F_B - F_A$ is the free-energy difference, $\beta = 1/kT$ is the reciprocal temperature in energy units, and W is the work involved in perturbing from system A into system B for a given configuration. In terms of the molecular potential energy U, $W = U_B - U_A$. Subscripts indicate values for the A and B systems, respectively. The FEP method is a special case of the more general nonequilibrium work (NEW) formalism of Jarzynski [6]. The NEW formalism expresses free-energy differences in terms of an average of the form of Eq. (1), but interpreting the average as one over an ensemble of initial conditions and permitting the transformation to be performed at any rate. FEP is an important special case because it can be implemented while sampling the "initial" system exclusively, so the free-energy calculation need not disturb the collection of other averages. Also, because FEP does not

require the lengthy calculations involved in the work process and reequilibrating afterwards, usually many more FEP measurements can be performed than is possible in the more general NEW case. On the other hand, to the extent that the NEW calculation is performed reversibly, fewer measurements are required to obtain a good estimate of the average. NEW measurements are also of interest because they can be performed on real systems in the laboratory [7].

Work distributions play a central role in the study and application of NEW calculations (including the FEP special case) [6,8]. The distribution $p_A(W)$ is the probability density for observing the work value W when performing the transformation $A \rightarrow B$ when beginning from an equilibrated A system. Likewise, $p_B(W)$ is the probability density for observing work $-W$ for a $B \rightarrow A$ transformation (we define W independently of the direction, so for FEP it is always $U_B - U_A$). The two conjugate work distributions are related [8,9]:

$$p_A(W)e^{-\beta W} = p_B(W)e^{-\beta \Delta F}. \tag{2}$$

The free-energy difference is given from the distributions via

$$\exp(-\beta \Delta F) = \int_{-\infty}^{\infty} p_A(W)e^{-\beta W} dW, \tag{3a}$$

$$\exp(+\beta \Delta F) = \int_{-\infty}^{\infty} p_A(W)e^{+\beta W} dW. \tag{3b}$$

A severe problem with the use of FEP calculations and NEW calculations, in general, is the strong tendency of the averages to exhibit bias or inaccuracy [10,11]. In some cases it is possible for repeated measurements to produce the same incorrect result, so it can be difficult to detect the inaccuracy from the simulation data alone. Moreover, the bias is often asymmetric in magnitude, meaning that a NEW calculation performed by averaging $A \rightarrow B$ work processes is different from the bias obtained using $B \rightarrow A$ processes [3,12–14].

In previous studies we have examined this bias, and most recently we have considered it for the special case in which the distribution of work values follows a Gaussian form [15]. However, the Gaussian-work model is incapable of characterizing asymmetric bias. Our interest in the present work is to introduce models that can capture this feature of NEW calculations. We are interested in particular in developing and examining models that are based on a Hamiltonian, rather than one that is defined directly in terms of the work distributions. A Hamiltonian-based model is defined on a phase space, and we think that basic understanding of the nature of the NEW bias can be advanced by looking at the systems from this perspective. Such insight can permit one to develop some intuition about the best way to perform a NEW calculation for a given physical system. However, it is very difficult to derive work distributions from a Hamiltonian model for a general NEW calculation, as this requires a detailed treatment of the dynamics of the work process [16]. Instead we focus on modeling the FEP calculation, for which we can derive the work distributions from the joint density of states of the A and B systems.

In the following section we review the neglected-tail model that we proposed in previous work as a means to estimate the bias from the work distributions. Then in Sec. III we present the two Hamiltonian-based models that we wish to use to study the asymmetric FEP bias. In Sec. IV we present results demonstrating that the neglected-tail model is effective in characterizing the bias for these models, and we examine the behavior of the bias as a function of the parameters of the models. We conclude in Sec. V.

II. NEGLECTED-TAIL BIAS MODEL

The neglected-tail model [14,15] can be understood with reference to Eqs. (3). For the $A \rightarrow B$ direction [Eq. (3a)], bias arises from poor sampling the negative tail of $p_A(W)$, which contributes greatly to the average because of its multiplication by the exponential $e^{-\beta W}$. Likewise, for the $B \rightarrow A$ direction [Eq. (3b)], the bias results from the failure to sample the upper values of $p_B(W)$. Clearly, if $p_A(W)$ and $p_B(W)$ are not symmetric (i.e., they differ by more than a simple translation in W), the bias will not be same for the two directions.

The neglected-tail model for the bias assumes that *all* of the error arises from failure to sample the relevant tail of the work distribution. In particular, we assume that the NEW process samples perfectly the work distribution for work values greater than (for $A \rightarrow B$) or less than (for $B \rightarrow A$) a limiting work value W^* and that it does not sample any work values at all beyond W^*. The likelihood that a NEW calculation involving M work measurements will have a limiting work value W^* is given by the probability distribution

$$P_A^*(W^*;M) = Mp_A(W^*)[1 - C_A(W^*)]^{M-1},$$

$$P_B^*(W^*;M) = Mp_B(W^*)[C_B(W^*)]^{M-1}, \tag{4}$$

where C_X is the cumulative distribution function

$$C_X(W^*) = \int_{-\infty}^{W^*} p_X(W), \quad X \in (A,B). \tag{5}$$

Upon increasing M, the peak in P_A^* will move to more negative W^* and the peak in P_B^* will move to more positive W^*. The bias in the free energy obtained when neglecting the tail beyond W^* is

$$\hat{\Delta F}_{A\rightarrow B}(W^*;M) - \Delta F = -kT \ln\left[\frac{1}{M}e^{-\beta(W^*-\Delta F)} + \left(1 - \frac{1}{M}\right)\frac{1 - C_B(W^*)}{1 - C_A(W^*)}\right], \tag{6a}$$

$$\hat{\Delta F}_{B\rightarrow A}(W^*;M) - \Delta F = +kT \ln\left[\frac{1}{M}e^{+\beta(W^*-\Delta F)} + \left(1 - \frac{1}{M}\right)\frac{C_A(W^*)}{C_B(W^*)}\right], \tag{6b}$$

which is derived using Eq. (2). The expected bias as a function of sampling length is obtained by averaging the biases given by Eq. (6) over the W^* values distributed according to Eqs. (4):

$$B_{A\to B}(M) = \int_{-\infty}^{\infty} dW^* P_A^*(W^*;M)[\hat{\Delta F}_{A\to B}(W^*;M) - \Delta F],$$

$$B_{B\to A}(M) = \int_{-\infty}^{\infty} dW^* P_B^*(W^*;M)[\hat{\Delta F}_{B\to A}(W^*;M) - \Delta F]. \tag{7}$$

There are a few approximations we can consider from this point. First we approximate Eq. (6) via

$$\hat{\Delta F}_{A\to B}(W^*;M) - \Delta F = -kT\ln[1 - C_B(W^*)],$$

$$\hat{\Delta F}_{B\to A}(W^*;M) - \Delta F = +kT\ln[C_A(W^*)]. \tag{8}$$

This form is appropriate for $M \gg 1$. As another approximation, we consider that the integral over the W^* distribution [Eq. (7)] may be approximated by the value of the bias at the mode or mean of $P^*(W^*)$.

Note that Eq. (8) indicates that the bias is characterized by the area of the conjugate distribution beyond W^* [e.g., for the $A\to B$ direction the bias is related to the area of $p_B(W)$ that is below W^*]. Thus, if the sampled distribution encompasses its conjugate, W^* is likely to lie beyond the bulk of the conjugate distribution and the bias will be small. This indicates that the bias should be less when sampling the broader distribution. We will refer to sampling in this direction as the "insertion" direction, while the direction that involves sampling the narrower $p(W)$ distribution we will refer to as "deletion." (These names derive from the application of FEP to calculate the chemical potential, which involves the insertion or deletion of a particle; usually the insertion direction is characterized by a broader distribution of work values.) In some cases the widths of the distributions are not so different such that one direction is clearly "insertion" and the other "deletion." For Gaussian work distributions, $p_A(W)$ and $p_B(W)$ are both Gaussians with equal variance [this equality is a consequence of Eq. (2)] and differ only by shifting in W [11,15]. In this case the bias is independent of the direction $A\to B$ versus $B\to A$, and the "insertion" and "deletion" labels are not meaningful.

III. TWO MODEL SYSTEMS

We now introduce the two model systems that are the focus of this work. Both models are formed from N noninteracting particles, with each particle under the influence of a common single-body potential that differs between the A and B systems. We can identify a phase (or configuration) space for these systems, and from this standpoint we can derive expressions for the work distributions they would exhibit in a FEP calculation. Although the behavior of each particle is independent of the others, the bias in the FEP calculation depends on their collective statistics. This combination of features of the model and the FEP calculation makes their analysis tractable but interesting and nontrivial.

A. Independent harmonic oscillators

The first model for consideration is a system of N independent particles under the influence of a harmonic potential of width and position that differ in the two systems: thus,

$$U_A = \sum_{i=1}^{N} \omega_A x_i^2,$$

$$U_B = \sum_{i=1}^{N} \omega_B (x_i - x_0)^2, \tag{9}$$

where x_i is a coordinate for particle i. The free-energy difference is $\Delta F = \frac{1}{2}NkT\ln(\omega_B/\omega_A)$. Increasing the parameter x_0 causes the important parts of phase space for the two systems to move apart. This makes sampling of one system less likely to encounter configurations important to the other, thereby increasing the bias symmetrically without affecting the true free-energy difference. Increasing the parameter ω_A or ω_B narrows the important phase space for the corresponding system. The more narrowly sampled system will be less likely to encounter configurations important to the other system, thus increasing the bias asymmetrically. Increasing N tends to exacerbate the bias tendencies imposed by the other parameters. Thus, by changing the parameters N, ω_A, ω_B, and x_0, we can change the degree of the asymmetry of the two work distributions $p_A(W)$ and $p_B(W)$ and construct different bias cases for study. In the Appendix we develop analytic expressions for the work distributions. The result for $p_A(W)$ is given in the Appendix [Eq. (A13)]. The distribution $p_B(W)$ can be obtained most conveniently via Eq. (2).

B. Two-state discrete model

Next, we consider a case with discrete work distributions. Here we divide the phase space into two parts, describing two possible states that a particle can occupy. We designate state 0 to form a fraction f of the total phase space, while state 1 occupies the fraction $1-f$. System A is defined such that a particle in state 0 has energy ε_{A0}, while in state 1 its energy is ε_{A1}. System B has energies ε_{B0} and ε_{B1} defined similarly. The total energy of a given configuration of a collection of N such particles in the A and B systems, respectively, is

$$U_A = n\varepsilon_{A0} + (N-n)\varepsilon_{A1},$$

$$U_B = n\varepsilon_{B0} + (N-n)\varepsilon_{B1}, \tag{10}$$

where n is the number of particles in state 0 in the given configuration. The partition function for each system is given as a sum over all values of n:

$$Q_X = \sum_{n=0}^{N} \binom{N}{n} f^n (1-f)^{N-n} \exp[-\beta U_X(n)] \quad [X \in (A,B)]$$
$$= [fe^{-\beta\varepsilon_{X0}} + (1-f)e^{-\beta\varepsilon_{X1}}]^N. \tag{11}$$

The free-energy difference is given by

$$\Delta F = -kTN \ln\left[\frac{fe^{-\beta\varepsilon_{B0}} + (1-f)e^{-\beta\varepsilon_{B1}}}{fe^{-\beta\varepsilon_{A0}} + (1-f)e^{-\beta\varepsilon_{A1}}}\right]. \tag{12}$$

This model could characterize a system of uncoupled molecules (or spins, perhaps) that can occupy either of two states (0 or 1), the properties of which are subject to the imposition of an external field (which is off in system *A* and on in system *B*, for example).

The work distributions for this discrete model are easily related to the *n* distributions, because the *n* value of a configuration is sufficient to specify *W* via

$$W(n) = U_B - U_A = n(\varepsilon_{B0} - \varepsilon_{A0} - \varepsilon_{B1} + \varepsilon_{A1}) + N(\varepsilon_{B1} - \varepsilon_{A1}). \tag{13}$$

The *n* values are distributed according to the binomial distribution

$$p_X(n) = \binom{N}{n} r_X^n (1-r_X)^{N-n}, \quad X \in (A,B), \tag{14}$$

where $r_X = fe^{-\beta\varepsilon_{X0}}[fe^{-\beta\varepsilon_{X0}} + (1-f)e^{-\beta\varepsilon_{X1}}]^{-1}$. Then $p_X(W) = p_X(n(W))|dn/dW|$ where $n(W)$ is the inverse of Eq. (13).

The tail distributions are similar to those of continuous tail distributions, except now we need to account for the possibility that the extreme value of *n*, labeled *n**, could be encountered more than once in a FEP calculation. Thus instead of Eq. (4), we have the probability that the smallest (largest) value of *n* is encountered exactly *m* times in a calculation:

$$P_A^*(n^*;m;M) = \binom{M}{m}[p_A(n^*)]^m[1 - C_A(n^*)]^{M-m},$$

$$P_B^*(n^*;m;M) = \binom{M}{m}[p_B(n^*)]^m[C_B(n^*-1)]^{M-m}, \tag{15}$$

where the cumulative distributions are $C_X(n^*) = \sum_{n=0}^{n^*} p_X(n)$ [the probability corresponding to Eq. (4) is the special case of $m=1$]. Correspondingly, the *m*-dependent free energies in two directions are [cf. Eq. (6)]

$$\hat{\Delta F}_{A\to B}(n^*;m;M) - \Delta F = -kT\ln\left[\frac{m}{M}e^{-\beta(W(n^*)-\Delta F)} + \left(1-\frac{m}{M}\right)\frac{1-C_B(n^*)}{1-C_A(n^*)}\right], \tag{16a}$$

$$\hat{\Delta F}_{B\to A}(n^*;m;M) - \Delta F = +kT\ln\left[\frac{m}{M}e^{+\beta(W(n^*)-\Delta F)} + \left(1-\frac{m}{M}\right)\frac{C_A(n^*-1)}{C_B(n^*-1)}\right]. \tag{16b}$$

Finally, the biases for the two directions according to this implementation of the neglected-tail model are

TABLE I. Summary of parameter sets for the harmonic-model systems examined in Figs. 1–4. In all cases, $\beta=1$ and $\omega_A=1$.

Case	N	$\beta\omega_A N x_0^2$	ω_B/ω_A	x_0	$\beta\Delta F$
A	20	0.002	2	0.01	6.93
B	20	20	2	1	6.93
C	20	0.002	10	0.01	23.0
D	50	0.005	2	0.01	17.3
E	50	50	2	1	17.3
F	50	0.005	10	0.01	57.6

$$B_{A\to B}(M) = \sum_{n^*=0}^{N}\sum_{m=1}^{M} P_A^*(n^*;m;M) \times[\hat{\Delta F}_{A\to B}(n^*;m;M) - \Delta F],$$

$$B_{B\to A}(M) = \sum_{n^*=0}^{N}\sum_{m=1}^{M} P_B^*(n^*;m;M) \times[\hat{\Delta F}_{B\to A}(n^*;m;M) - \Delta F]. \tag{17}$$

For large *M* it is expensive to perform the full sum over *m*. Fortunately, in these cases the *m* dependence of the probabilities is sharply peaked, so we may use just a few terms about the maximum. The *m* where *P** has its maximum is in the vicinity of

$$m_{\max}^{A\to B}(n^*) \approx \frac{Mp_A(n^*) - [1 - C_A(n^*)]}{[1 - C_A(n^*)] + p_A(n^*)},$$

$$m_{\max}^{B\to A}(n^*) \approx \frac{Mp_B(n^*) - C_B(n^*-1)}{C_B(n^*-1) + p_B(n^*)}. \tag{18}$$

IV. RESULTS AND DISCUSSION

Let us consider first the model of uncoupled harmonic particles. Without loss of generality we choose $\omega_B > \omega_A$; consequently, $p_A(W)$ is always wider than $p_B(W)$ and $A\to B$ always represents the "insertion" direction. We have examined the bias for several values of the model parameters, as summarized in Table I. The bias as calculated by the neglected-tail model for these cases is presented in Fig. 1 as a function of sampling amount *M*. Also shown in the figure are results of numerical calculations performed to test the validity of the bias model. Here we conducted simulations in which we sampled the configurations of the particles, sampling directly from a Gaussian distribution to get a Boltzmann sample of configurations for the given system (*A* or *B*). For each configuration we evaluated the work in perturbing to the other system ($A\to B$ or $B\to A$) and evaluated the free-energy difference for *M* such realizations (up to $M=10^6$). This process was repeated approximately 5000–12 000 times (depending on *N*), and the average bias in the free energy over all *M*-length samples was evaluated. These results are presented

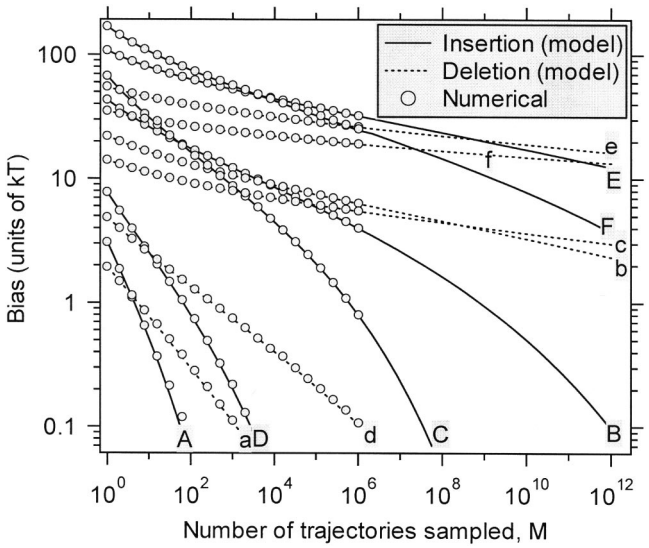

FIG. 1. Bias (in absolute value) in the free-energy difference for the harmonic model, as a function of the number of free-energy perturbations performed. Circles describe the results of numerical calculations (which were performed only up to $M=10^6$), and lines are given by the neglected-tail bias model. Letters describe different model systems as keyed in Table I. Uppercase letters (and solid lines) are for the insertion direction ($A\rightarrow B$), and lowercase letters (dashed lines) are for the deletion direction ($B\rightarrow A$).

as a function of M with the model data. The model agrees perfectly with the numerical results for the range of M tested, and it is reasonable to expect that the agreement would continue into the range of larger M shown in the figure. Discrepancies arise only for small values of the bias (below the scale of the figure), which is the range in which the variance of the data is greater than the (square) error due to the bias. Thus the model is applicable in the cases where the primary source of error is the bias rather than the noise. It should be noted that for the deletion calculations the plot presents the negative of the bias. It is well known that the biases in the two directions are of opposite sign, and the present results are consistent with this expectation.

Figure 2 shows the insertion and deletion work-distribution pairs for case C in Table I, together with their

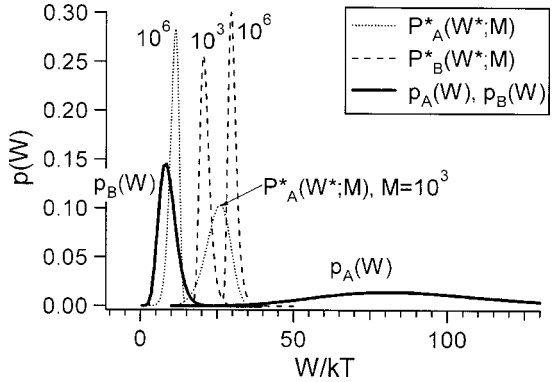

FIG. 2. Work distributions for case C defined in Table I. Also shown are the extreme-work distributions as given by Eq. (4) for two values of M.

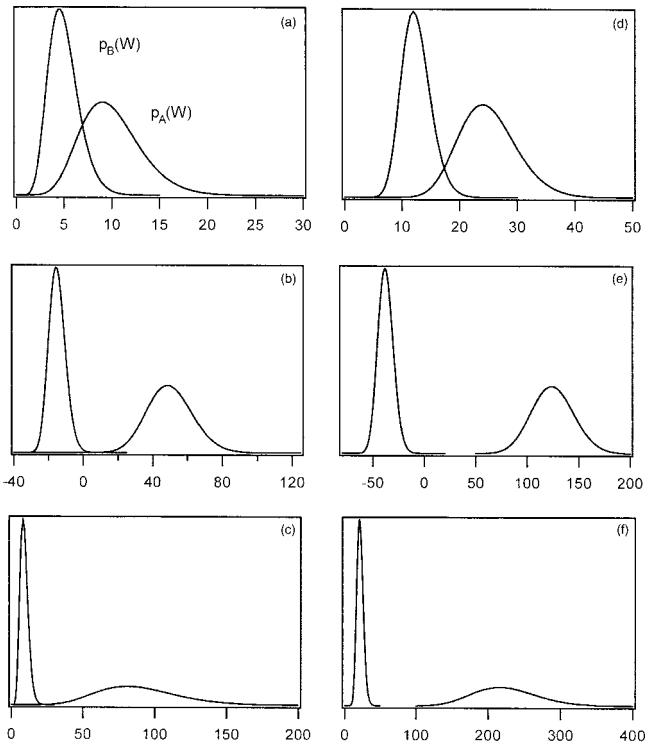

FIG. 3. Work distributions for different harmonic-model systems as keyed in Table II and with FEP bias as shown in Fig. 1. The abscissa is the work (in units of kT). The rightmost distribution is $p_A(W)$, and the leftmost distribution is $p_B(W)$. All curves are normalized to unity.

tail-sampling distributions according to Eq. (4) at $M=10^3$ and $M=10^6$. At $M=10^6$, W^* in $A\rightarrow B$ (insertion) direction is already in the middle of $p_B(W)$, while W^* in the $B\rightarrow A$ (deletion) direction is still on the near edge of $p_A(W)$. The neglected-tail model indicates that the W^* distribution must peak on the other side of the complementary work distribution for the bias to be small. One can see how this happens first for the insertion direction as M is increased. Figure 3 presents work distributions for the different cases presented in Table I and Fig. 1. The work distribution $p_A(W)$ is always to the right simply because we choose to define W as U_B-U_A; it is always the wider distribution because we select the system A such that $\omega_B \geqslant \omega_A$.

Considering now the behavior of the bias, we see in Fig. 1 that the bias in the insertion direction decreases much more quickly than the bias in the deletion direction. This supports our contention that better results are obtained when sampling the system having the wider work distribution. Comparison with Fig. 3 shows that the asymmetry in the bias connects to the difference in the width of the work distributions and (unsurprisingly) that the bias decreases more slowly as the distributions become farther apart. In some situations the bias in the deletion direction shows little sign of diminishing even for very large amounts of sampling, while for the same case the insertion bias becomes small after comparatively little sampling. Case C is such an example. It can be noted that for this pair of systems, x_0 is nearly zero and ω_B/ω_A is large. Thus the important phase space of the B system is smaller

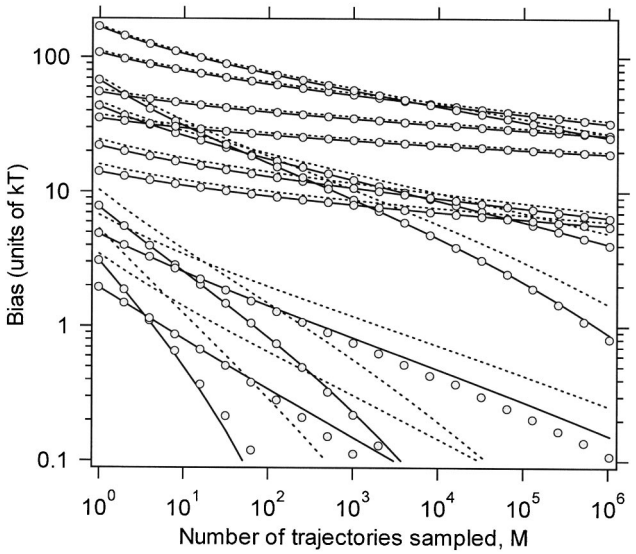

FIG. 4. Demonstration of approximations applied to the neglected-tail model. Circles are the numerical calculations as presented in Fig. 1 (including both insertion and deletion). Solid lines use Eq. (6), and dashed lines use Eq. (8), both with W^* given by the mean of $P_X^*(W^*)$ [instead of integrating over W^* as indicated by Eq. (7)].

than that of the A system, but it is not displaced from it. This is an example of the "subset" phase-space relation discussed in previous work [3,12]. In this situation a single-stage FEP calculation can provide a useful result, but only if applied in the right direction. On the other hand, in some cases the bias in both directions is slow to decay (although insertion still has advantage) and a single-stage calculation cannot provide a good result for any feasible amount of sampling. Cases E and F provide examples of this situation. In case E, x_0 is large, and any phase-space overlap is not enough to form a subset relation; in case F, x_0 is small, but N is large and the important phase space of B is much too small a part of A to permit it to be sampled well in an $A \rightarrow B$ perturbation. In these cases a multistage approach would be needed to obtain good results.

We would also like to know the performance of the approximated bias model by using Eq. (6) or Eq. (8) with W^* to be the mean or mode of P^* distributions, because these approximations simplify the calculations considerably. In Fig. 4 we make these comparisons, which show that using Eq. (6) with the mean value of P^* is good enough to approximate the biases for this harmonic model, while Eq. (8) works well only for the large bias values. Also it is better to use the mean value of P^* instead of the mode value, because the P^* distribution itself is asymmetric (it has a longer tail on one side than the other side), so the mean is more representative than the mode.

The neglected-tail model and Eqs. (4) and (6), in particular, indicate that the bias is entirely determined (at least at large M) from the work distributions $p_X(W)$, either directly or through the cumulative distributions $C_X(W)$. Thus we can consider the behavior in terms of the groups that appear in the expressions derived above for the work distributions. For the harmonic model, if we consider the bias scaled by the temperature (βW), then the only independent groups appearing in Eq. (A13) for $p_A(W)$ are $\beta\omega_A N x_0^2$, ω_B/ω_A, and N. Figure 5 provides a parametric study of the shape of the work distributions in terms of these groups (see Table II). The overlap between the distributions decreases as any of these parameters are increased, but each has a different effect on the nature of the phase-space overlap of the A and B systems [3,12]. The group $\beta\omega_A N x_0^2$ promotes separation of the important phase spaces, while the ratio ω_B/ω_A promotes the diminishing of the B-important space relative to that for A, keeping the separation fixed. From the standpoint of the work distributions alone this distinction is difficult to make. The cases shown in the figure cover a wide range of behaviors of the work distributions, and their connection to the Hamiltonian-based model may prove useful in interpreting them. Such insight can guide the formulation of multistage NEW methods for application in a given situation.

We turn now to the two-state discrete model, with parameters chosen as summarized in Table III. For these cases Fig. 6 presents the bias as a function of sampling length, according to both the neglected-tail model and numerical calculations of the type used in Fig. 1 for the harmonic model (but now using 20 000 – 40 000 free-energy calculations to compute the average bias). Again the agreement is excellent for the bias values shown in the figure, which represents the range where the bias is more of a problem than the stochastic error (noise). As in the case of the harmonic model, the bias in the insertion direction decays significantly more quickly than that in the deletion direction. One might note that in both models, at small M, the deletion bias is often smaller than that for insertion. In these cases, however, the bias is still quite large, and the results from neither direction could be considered useful. The faster decay of the insertion bias is a much more useful feature, as it permits the insertion calculation to yield a useful result with considerably smaller sampling amounts. Work distributions are presented in Fig. 7 for some of the systems described. The relation between the work distributions and the bias is qualitatively similar to that observed for the harmonic model.

V. CONCLUSIONS

We have developed formulas for the work distributions for two types of Hamiltonian-based model systems. We have applied the neglected-tail bias model for nonequilibrium work calculations, and through numerical studies we have demonstrated that the model is effective in characterizing the bias in these systems. We have shown that the models exhibit asymmetric bias, in that the magnitude of the bias is larger when the perturbation is applied in one direction versus the other. The neglected-tail model indicates that the smaller bias is obtained when sampling the system with the broader work distribution, perturbing into the system with the narrower distribution.

One of our interests in presenting and examining these models is in helping us to understand the relation between the performance of a NEW or FEP calculation and the nature of the overlap of the important phase spaces of the two sys-

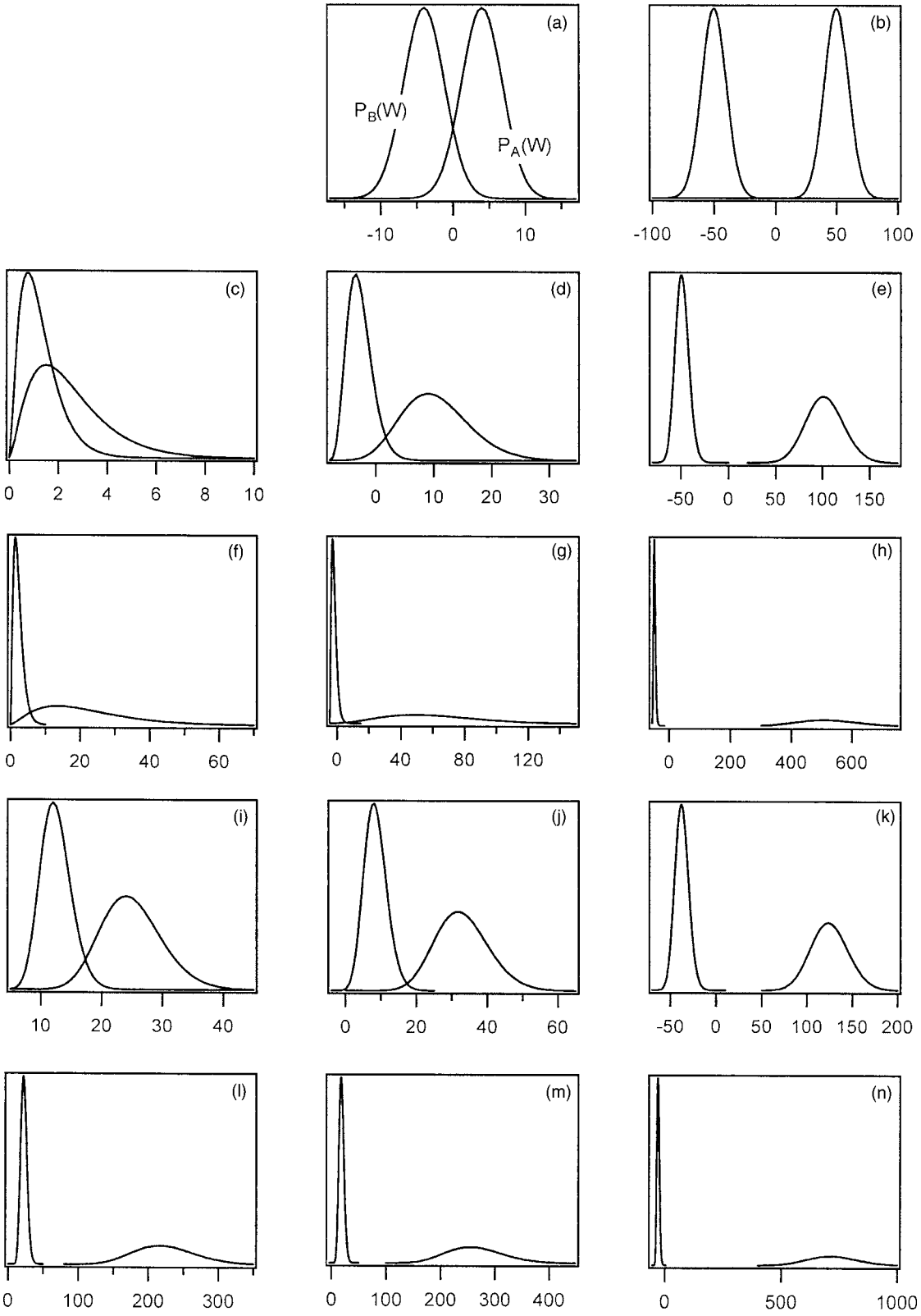

FIG. 5. Work distributions for different harmonic-model systems as keyed in Table II. The abscissa is the work (in units of kT). The rightmost distribution is $p_A(W)$, and the leftmost distribution is $p_B(W)$. All curves are normalized to unity. Figures are arranged such that those more to the right represent larger $\beta\omega_A N x_0^2$ and those further down (for a given N) represent larger ω_B/ω_A. First three rows are $N=5$, and last two rows are $N=50$.

TABLE II. Summary of parameter sets for the harmonic-model systems examined in Fig. 5. In all cases $\beta=1$ and $\omega_A=1$.

Case	N	$\beta\omega_A N x_0^2$	ω_B/ω_A	x_0	ΔF
A	5	4	1	0.894	0
B	5	50	1	3.16	0
C	5	0	2	0	1.73
D	5	4	2	0.894	1.73
E	5	50	2	3.16	1.73
F	5	0	10	0	5.76
G	5	4	10	0.894	5.76
H	5	50	10	3.16	5.76
I	50	0	2	0	17.3
J	50	4	2	0.283	17.3
K	50	50	2	1	17.3
L	50	0	10	0	57.6
M	50	4	10	0.283	57.6
N	50	50	10	1	57.6

tems of interest. The models presented here are defined in terms of particles sampling a phase space, so they can be studied from this perspective. We think this point of view can be helpful in choosing and formulating multistage methods that can greatly aid in obtaining accurate free-energy differences for a minimum amount of sampling. Investigations in this direction are currently underway.

ACKNOWLEDGMENTS

This work is supported by the Division of Chemical Sciences, Office of Basic Energy Sciences of the U.S. Department of Energy (Contract No. DE-FG02-96ER14677). Computer resources have been provided by the University at Buffalo Center for Computational Research.

APPENDIX

Here we derive the instantaneous-change work distributions for the system of uncoupled harmonic oscillators. We

TABLE III. Summary of parameter sets for the discrete-model systems examined in Figs. 6 and 7. In all cases $\beta=1$ and $\varepsilon_{A0}=\varepsilon_{B1}=0$.

Case	N	f	ε_{A1}	ε_{B0}	ΔF
A	50	0.3	0.3	0.5	−3.73
B	50	0.3	0	2	15.0
C	50	0.3	0	5	17.7
D	50	0.3	3	5	−37.0
E	50	0.9	0.3	0.5	20.5
F	50	0.9	0	2	75.3
G	20	0.3	0	2	6.01
H	20	0.3	0	5	7.08
I	20	0.4	0.1	50	9.04

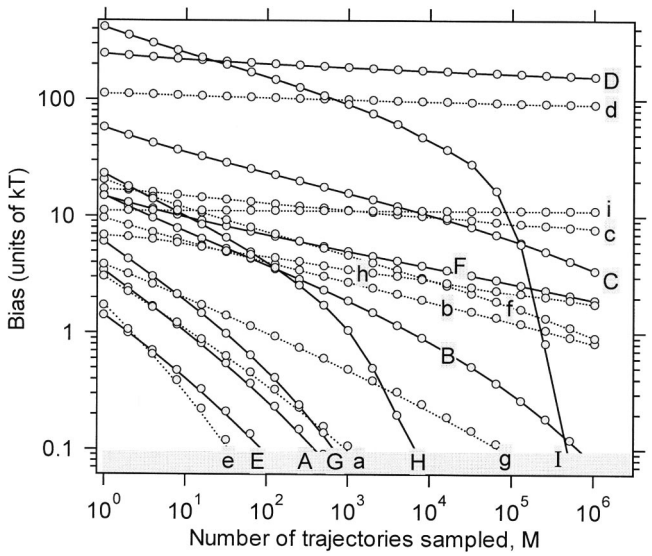

FIG. 6. Bias (in absolute value) in the free-energy difference for the discrete model, as a function of the number of free-energy perturbations performed. Circles describe the results of numerical calculations, and lines are given by the neglected-tail bias model. Letters describe different model systems as keyed in Table III. Uppercase letters (and solid lines) are for the insertion direction ($A\rightarrow B$), and lowercase letters (dashed lines) are for the deletion direction ($B\rightarrow A$).

begin by deriving the joint density of states for the energies of the A and B systems. Thus we define $\Omega(U_A,U_B)$ as the probability density for configurations that have energy U_A in the A system and at the same time have energy U_B in the B system. This distribution can be represented via integrals of δ functions:

$$\Omega(U_A,U_B)=\int_N dx^N \delta\left(U_A-\omega_A\sum_{i=1}^{N}x_i^2\right)\times\delta\left(U_B-\omega_B\sum_{i=1}^{N}(x_i-x_0)^2\right). \tag{A1}$$

As a notational convenience we introduce the scaled energies

$$\tilde{U}_A=\frac{U_A}{\omega_A N x_0^2},\quad \tilde{U}_B=\frac{U_B}{\omega_B N x_0^2}, \tag{A2}$$

then

$$\Omega(U_A,U_B)=(\omega_A\omega_B)^{-1}(Nx_0^2)^{N/2-2}\int_N d\tilde{x}^N\delta\left(\tilde{U}_A-\sum_{i=1}^{N}\tilde{x}_i^2\right)\times\delta\left(\tilde{U}_B-\sum_{i=1}^{N}(\tilde{x}_i-N^{-1/2})^2\right). \tag{A3}$$

The δ functions are implemented using the integral identity

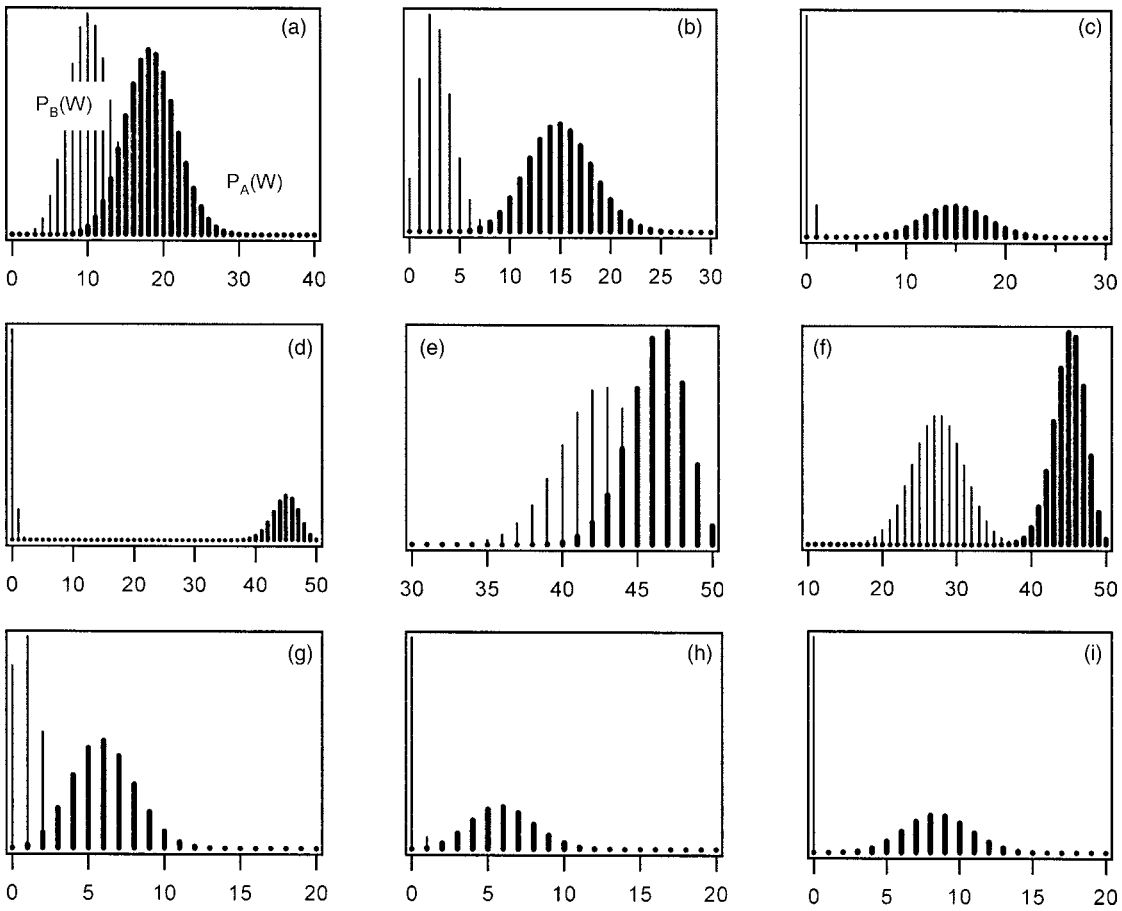

FIG. 7. Work distributions for different discrete-model systems as keyed in Table III and with FEP bias as shown in Fig. 6. The abscissa is the work (in units of kT). The rightmost distribution (using thick lines) is $p_A(W)$, and the leftmost distribution (thin lines) is $p_B(W)$. All distributions are normalized to unity.

$$\delta(x) = \frac{1}{2\pi i}\int_C dz\, \exp(xz), \tag{A4}$$

where the integral is performed along a vertical contour in the complex plane, with an infinitesimal but nonzero positive real component. Substituting in Eq. (A3),

$$\Omega(U_A,U_B) = \frac{(\omega_A\omega_B)^{-1}(Nx_0^2)^{N/2-2}}{(2\pi i)^2} \times \int_N d\tilde{x}^N \int_{C_A} dz_A \exp\left[z_A\left(\tilde{U}_A - \sum \tilde{x}_i^2\right)\right] \times \int_{C_B} dz_B \exp\left[z_B\left(\tilde{U}_B - \sum (\tilde{x}_i - N^{-1/2})^2\right)\right]. \tag{A5}$$

We now exchange the order of integration and complete the integrals over the coordinates. Then,

$$\Omega(U_A,U_B) = (\omega_A\omega_B)^{-1}(Nx_0^2)^{N/2-2}\frac{\pi^{N/2}}{(2\pi i)^2} \times \int_{C_A} dz_A \int_{C_B} dz_B \exp(z_A\tilde{U}_A) \times \exp(z_B\tilde{U}_B)\exp\left(-\frac{z_A z_B}{z_A+z_B}\right)(z_A+z_B)^{-N/2}. \tag{A6}$$

We define z such that $z_B = z - z_A$ and complete the integral over z_A:

$$\Omega(U_A,U_B) = (\omega_A\omega_B)^{-1}(Nx_0^2)^{N/2-2}\frac{\pi^{(N+1)/2}}{(2\pi)^2 i} \times \int_C dz \exp\left[\frac{z}{4}[\tilde{U}_{A,\max}(\tilde{U}_B) - \tilde{U}_A] \times [\tilde{U}_A - \tilde{U}_{A,\min}(\tilde{U}_B)]\right] z^{-(N-1)/2}, \tag{A7}$$

where

$$\tilde{U}_{A,\min}(\tilde{U}_B) \equiv (\tilde{U}_B^{1/2} - 1)^2,$$

$$\tilde{U}_{A,\max}(\tilde{U}_B) \equiv (\tilde{U}_B^{1/2} + 1)^2, \tag{A8}$$

which yields

$$\Omega(U_A, U_B) = (\omega_A \omega_B)^{-1} (N x_0^2)^{N/2-2} \frac{\pi^{(N-1)/2}}{2\Gamma((N-1)/2)} \times \left\{ \frac{1}{4} [\tilde{U}_{A,\max}(\tilde{U}_B) - \tilde{U}_A][\tilde{U}_A - \tilde{U}_{A,\min}(\tilde{U}_B)] \right\}^{(N-3)/2}. \tag{A9}$$

The univariate density of states can be recovered by integration over $\tilde{U}_A$:

$$\Omega(U_B) = N x_0^2 \omega_A \int_{\tilde{U}_{A,\min}}^{\tilde{U}_{A,\max}} \Omega(\tilde{U}_A, U_B) d\tilde{U}_A = \frac{(\pi/\omega_B)^{N/2}}{\Gamma(N/2)} U_B^{N/2-1}, \tag{A10}$$

which is consistent with the previously established formula for the density of states of a N uncoupled harmonic oscillators [17].

To obtain the work distribution that would be observed when sampling the A ensemble, we need instead the joint density of states for U_A and $W = U_B - U_A$. This is easily obtained through manipulation of Eq. (A9), yielding

$$\Omega_W(U_A, W) = (\omega_A \omega_B)^{-1} (N x_0^2)^{N/2-2} \frac{\pi^{(N-1)/2}}{2\Gamma((N-1)/2)} \times \left[\frac{1}{4} \left(1 - \frac{\omega_A}{\omega_B}\right)^2 [\tilde{U}_{A,\max}(W) - \tilde{U}_A] \times [\tilde{U}_A - \tilde{U}_{A,\min}(W)] \right]^{(N-3)/2}, \tag{A11}$$

where

$$\tilde{U}_{A,\min}(W) \equiv \left(\frac{1 - \sqrt{D(W)}}{(1 - \omega_A/\omega_B)} \right)^2, \tag{A12a}$$

$$\tilde{U}_{A,\max}(W) \equiv \left(\frac{1 + \sqrt{D(W)}}{(1 - \omega_A/\omega_B)} \right)^2, \tag{A12b}$$

$$D(W) \equiv \frac{\omega_A}{\omega_B} + \frac{W}{\omega_B N x_0^2} \left(1 - \frac{\omega_A}{\omega_B}\right). \tag{A12c}$$

Then upon integration over the Boltzmann-weighted energies of system A and normalization, we obtain

$$p_A(W) = \int_{U_{A,\min}(W)}^{U_{A,\max}(W)} \Omega_W(U_A, W) e^{-\beta U_A} dU_A \Bigg/ \int_{W_{\min}}^{\infty} dW \int_{U_{A,\min}(W)}^{U_{A,\max}(W)} \Omega_W(U_A, W) e^{-\beta U_A} dU_A = \frac{\beta\omega_A/\omega_B}{1 - \omega_A/\omega_B} [D(W)]^{(N-2)/4} \exp\left(- \frac{\beta\omega_A N x_0^2 [1 + D(W)]}{(1 - \omega_A/\omega_B)^2} \right) I_{N/2-1}\left(\frac{2\beta\omega_A N x_0^2 \sqrt{D(W)}}{(1 - \omega_A/\omega_B)^2} \right), \tag{A13}$$

where $I_k(z)$ is the Bessel function and $D(W)$ is as defined in Eq. (A12c); the cumulative distributions are determined by numerical integration. Two special cases arise: when $x_0 = 0$ and when $\omega_A = \omega_B$, respectively. When $x_0 = 0$, the energy of one system uniquely determines the energy of the other, and U_A and U_B are not independent in the density of states. Thus

$$\Omega(U_A, U_B) = \frac{1}{\omega_A \omega_B} \frac{\pi^{N/2}}{\Gamma(N/2)} \left(\frac{U_B}{\omega_B} \right)^{N/2-1} \delta\left(\frac{U_A}{\omega_A} - \frac{U_B}{\omega_B} \right) \quad (x_0 = 0), \tag{A14}$$

and for this case

$$p_A(W) = \beta \left(\frac{\omega_A/\omega_B}{1 - \omega_A/\omega_B} \right)^{N/2} \frac{(\beta W)^{N/2-1}}{\Gamma(N/2)} \exp\left(- \frac{\omega_A/\omega_B}{1 - \omega_A/\omega_B} \beta W \right),$$

$$C_A(W) = 1 - \frac{\gamma\left(\frac{N}{2}, \frac{\omega_A/\omega_B}{1 - \omega_A/\omega_B} W \right)}{\Gamma\left(\frac{N}{2} \right)}, \tag{A15}$$

where $\gamma(n, z)$ is the incomplete gamma function. For the other special case, the work distribution is Gaussian:

$$p_A(W) = \left(\frac{\beta}{4\pi N x_0^2 \omega} \right)^{1/2} \exp\left[- \beta \frac{(W - \omega N x_0^2)^2}{4\omega N x_0^2} \right] \quad (\omega \equiv \omega_A = \omega_B). \tag{A16}$$

This case was treated in some detail in previous work [15].

[1] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. (Academic, San Diego, 2002).
[2] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed. (Wiley, New York, 1985).
[3] D. A. Kofke, Fluid Phase Equilib. **121**, 8742 (2004).
[4] R. J. Gdanitz, Curr. Opin. Solid State Mater. Sci. **3**, 414 (1998).
[5] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954).
[6] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997); Phys. Rev. E **56**, 5018 (1997).
[7] G. Hummer and A. Szabo, Proc. Natl. Acad. Sci. U.S.A. **98**, 3658 (2001); J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, and C. Bustamante, Science **296**, 1832 (2002); Biophys. J. **82**, 193a (2002).
[8] C. H. Bennett, J. Comput. Phys. **22**, 245 (1976).
[9] G. E. Crooks, Phys. Rev. E **61**, 2361 (2000).
[10] R. H. Wood, W. C. F. Mühlbauer, and P. T. Thompson, J. Phys. Chem. **95**, 6670 (1991); N. D. Lu, J. Adhikari, and D. A. Kofke, Phys. Rev. E **68**, 026122 (2003); D. M. Zuckerman and T. B. Woolf, Phys. Rev. Lett. **89**, 180602 (2002).
[11] J. Gore, F. Ritort, and C. Bustamante, Proc. Natl. Acad. Sci. U.S.A. **100**, 12 564 (2003).
[12] D. A. Kofke and P. T. Cummings, Mol. Phys. **92**, 973 (1997); D. A. Kofke, *ibid.* **102**, 405 (2004).
[13] N. D. Lu and D. A. Kofke, AIChE Symp. Ser. **97**, 146 (2001); J. Chem. Phys. **115**, 6866 (2001).
[14] N. D. Lu and D. A. Kofke, J. Chem. Phys. **114**, 7303 (2001).
[15] D. Wu and D. A. Kofke, J. Chem. Phys. **121**, 8742 (2004).
[16] O. Mazonka and C. Jarzynski, e-print cond-mat/9912121.
[17] C. Predescu, M. Predescu, and C. V. Ciobanu, J. Chem. Phys. **120**, 4119 (2004).