

Statistically optimal analysis of samples from multiple equilibrium states

Michael R. Shirts^{1,*} and John D. Chodera^{2,†}

¹*Department of Chemistry, Columbia University, New York, NY 10027*

²*Department of Chemistry, Stanford University, Stanford, CA 94305*

(Dated: June 17, 2008)

We present a new estimator for computing free energy differences and thermodynamic expectations as well as their uncertainties from samples obtained from multiple equilibrium states via either simulation or experiment. The estimator, which we term the multistate Bennett acceptance ratio (MBAR) estimator because it reduces to the Bennett acceptance ratio when only two states are considered, has significant advantages over multiple histogram reweighting methods for combining data from multiple states. It does not require the sampled energy range to be discretized to produce histograms, eliminating bias due to energy binning and significantly reducing the time complexity of computing a solution to the estimating equations in many cases. Additionally, an estimate of the statistical uncertainty is provided for all estimated quantities. In the large sample limit, MBAR is unbiased and has the lowest variance of any known estimator for making use of equilibrium data collected from multiple states. We illustrate this method by producing a highly precise estimate of the potential of mean force for a DNA hairpin system, combining data from multiple optical tweezer measurements under constant force bias.

I. INTRODUCTION

A recurring challenge in statistical physics, computational chemistry, and single molecule experiments is the collection of a sufficient amount of data to estimate physical quantities of interest to adequate precision. In computer simulations of physical or chemical models, such quantities include potentials of mean force, phase coexistence curves, fluctuation or temperature-dependent properties, and free energy differences. In single-molecule experiments, these quantities might include potentials of mean force along a pulling coordinate or the distance between fluorescence probes during resonant energy transfer. For all of these problems, collection of sufficient statistics for a reliable estimate often requires multiple simulations at different thermodynamic states [1] or measurements performed under different applied biasing potentials. In computer simulations, multi-state techniques such as umbrella sampling [2], simulated [3] and parallel tempering [4] and the use of alchemical intermediates in free energy calculations can greatly aid convergence; in experiments, data collected under constant applied force can help provide adequate sampling of conformations of interest [24].

Even with these methods, it may require a large quantity of data to produce estimates with the desired precision. Computing the most precise estimate possible from the available data can therefore be critical in allowing these quantities to be estimated with reasonable computational or experimental effort. While the choice of thermodynamic states to sample can also greatly affect efficiency, we focus here on only the problem of sta-

tistically efficient estimation given samples from predetermined states.

Early methods for computing free energy differences [5, 6] or equilibrium expectations [2] relied upon one-sided exponential averaging (EXP), which is formally exact but does not make the most efficient use of data when samples from more than one state are available [7]. Subsequently, the Bennett acceptance ratio method (BAR) [8, 9] greatly improved upon EXP for the computation of free energy differences, producing statistically optimal estimates of free energy differences when two states are sampled [9] and yielding estimates that can be more than an order of magnitude more precise [7]. More recently, multiple histogram reweighting methods [10, 11] were proposed as a way to incorporate data from multiple states to produce superior estimates of free energy differences and equilibrium expectations for arbitrary thermodynamic states, including states not sampled.

While multiple histogram techniques — most notably, the weighted histogram analysis method (WHAM) [11] — can produce statistically optimal estimates of the discretized densities of states [10] or histogram occupation probabilities [12], they have several limitations for the treatment of continuous systems. First, the reliance on energy histograms of width sufficient to contain many samples — often larger than many times the thermal energy — introduces a bias that can be substantial and often difficult to assess [13]. Second, unlike BAR, there are no direct expressions to estimate the statistical uncertainty in free energy differences or expectations obtained from WHAM. Third, application of WHAM to a samples collected with a biasing potential that is not trivially scaled by a linear field parameter requires a number of bins that grows exponentially in the number of states, making it computationally intractable for even modest numbers of states. While more recent maximum likelihood [12] and Bayesian formulations [14] mitigate

*Electronic address: michael.shirts@columbia.edu

†Electronic address: jchodera@stanford.edu

the memory requirements, they do not remove the histogram bias effects, and introduce a costly Markov chain Monte Carlo sampling procedure to estimate uncertainties [14, 15].

Here, we use recent results from the field of statistical inference [16, 17, 18, 19] to construct a statistically optimal estimator for computing free energy differences and equilibrium expectations at arbitrary thermodynamic states, using equilibrium samples from multiple thermodynamic states. The resulting estimator, termed the multistate Bennett acceptance ratio (MBAR) estimator as it reduces to Bennett’s method when only two states are considered [8], is equivalent to WHAM in the limit that histogram bin widths are shrunk to zero but is derived without the need to invoke histograms. Unlike WHAM, this estimator provides a direct assessment of uncertainties, critical in making comparison between experiment and theory, and the computational expense of computing the estimator remains modest across a wider variety of applications. Furthermore, it can easily be applied to data sampled from non-Boltzmann sampling schemes, or to the analysis of single-molecule experiments in cases where an external bias potential is applied.

This paper is organized as follows: Section II recapitulates the literature on extended bridge sampling estimators, used here as the basis for the MBAR estimator. Expressions for computing estimates of free energy differences (Section III) and equilibrium expectations (Section IV) are then provided. Finally, we illustrate the method in Section V by applying it to the estimation of the potential of mean force for a DNA hairpin system by combining data from multiple equilibrium optical force clamp experiments under different external biasing potentials.

II. EXTENDED BRIDGE SAMPLING ESTIMATORS

Suppose we obtain N_i uncorrelated equilibrium samples from each of K thermodynamic states within the same ensemble, such as NVT, NPT, or μ VT (see Appendix A for more information on subsampling correlated timeseries data to produce uncorrelated samples). Each state is characterized by a specified combination of inverse temperature, potential energy function, pressure, and/or chemical potential(s), depending upon the ensemble. We define the *reduced potential function* $u_i(\mathbf{x})$ for state i to be

$$u_i(\mathbf{x}) = \beta_i[U_i(\mathbf{x}) + p_i V(\mathbf{x}) + \boldsymbol{\mu}_i^T \mathbf{n}(\mathbf{x})] \quad (1)$$

where $\mathbf{x} \in \Gamma$ denotes the configuration of the system within a configuration space Γ , with volume $V(\mathbf{x})$ (in the case of a constant pressure ensemble) and $\mathbf{n}(\mathbf{x})$ the number of molecules of each of M components of the system (in the case of a (semi)grand ensemble). For each state i , β_i denotes the inverse temperature, $U_i(\mathbf{x})$ the potential energy function (which may include biasing

weights), p_i the external pressure, and $\boldsymbol{\mu}_i$ the vector of chemical potentials of the M system components.

Configurations $\{\mathbf{x}_{in}\}_{n=1}^{N_i}$ from state i are sampled from the probability distribution

$$p_i(\mathbf{x}) = c_i^{-1} q_i(\mathbf{x}) \quad ; \quad c_i = \int_{\Gamma} d\mathbf{x} q_i(\mathbf{x}) \quad (2)$$

where $q_i(\mathbf{x})$ is here nonnegative and represents an unnormalized density function, and c_i is the (generally unknown) normalization constant (known in statistical mechanics as the *partition function*). In samples obtained from standard Metropolis Monte Carlo or molecular dynamics simulations or from experiment, this unnormalized density is simply the Boltzmann weight $q_i(\mathbf{x}) = \exp[-u_i(\mathbf{x})]$ but may in general differ in simulations employing non-Boltzmann weights, such as multicanonical simulations [20] and those using Tsallis statistics [21].

We wish to produce an estimator for the difference in dimensionless free energies

$$\Delta f_{ij} \equiv f_j - f_i = -\ln \frac{c_j}{c_i} = -\ln \frac{\int_{\Gamma} d\mathbf{x} q_j(\mathbf{x})}{\int_{\Gamma} d\mathbf{x} q_i(\mathbf{x})} \quad (3)$$

where the f_i are related to the dimensional free energies F_i by $f_i = \beta_i F_i$, and also the equilibrium expectations

$$\langle A \rangle_i \equiv \int_{\Gamma} d\mathbf{x} p_i(\mathbf{x}) A(\mathbf{x}) = \frac{\int_{\Gamma} d\mathbf{x} A(\mathbf{x}) q_i(\mathbf{x})}{\int_{\Gamma} d\mathbf{x} q_i(\mathbf{x})}. \quad (4)$$

These expectations can be computed as ratios of the normalization constants if we define new functions $q(\mathbf{x}) = A(\mathbf{x}) q_i(\mathbf{x})$, where the $q(\mathbf{x})$ no longer need be nonnegative for states from which no samples are collected [22].

To construct an estimator for these ratios of normalization constants, we first note the identity

$$\begin{aligned} c_i \langle \alpha_{ij} q_j \rangle_i &= \left[\int_{\Gamma} d\mathbf{x} q_i(\mathbf{x}) \right] \cdot \frac{\int_{\Gamma} d\mathbf{x} q_i(\mathbf{x}) \alpha_{ij}(\mathbf{x}) q_j(\mathbf{x})}{\int_{\Gamma} d\mathbf{x} q_i(\mathbf{x})} \\ &= \int_{\Gamma} d\mathbf{x} q_i(\mathbf{x}) \alpha_{ij}(\mathbf{x}) q_j(\mathbf{x}) \\ &= \left[\int_{\Gamma} d\mathbf{x} q_j(\mathbf{x}) \right] \cdot \frac{\int_{\Gamma} d\mathbf{x} q_j(\mathbf{x}) \alpha_{ij}(\mathbf{x}) q_i(\mathbf{x})}{\int_{\Gamma} d\mathbf{x} q_j(\mathbf{x})} \\ &= c_j \langle \alpha_{ij} q_i \rangle_j \end{aligned} \quad (5)$$

which holds for arbitrary choice of functions $\alpha_{ij}(\mathbf{x})$, provided the c_i are nonzero.

Using this relation, summing over the index j , and substituting the empirical estimator $N_i^{-1} \sum_{n=1}^{N_i} g(\mathbf{x}_{in})$ for the expectations $\langle g \rangle_i$, we obtain a set of K estimating equations

$$\sum_{j=1}^K \frac{\hat{c}_i}{N_i} \sum_{n=1}^{N_i} \alpha_{ij} q_j(\mathbf{x}_{in}) = \sum_{j=1}^K \frac{\hat{c}_j}{N_j} \sum_{n=1}^{N_j} \alpha_{ij} q_i(\mathbf{x}_{jn}) \quad (6)$$

for $i = 1, 2, \dots, K$, where solution of the set of equations for the \hat{c}_i yields an estimate of the c_i from the sampled data determined up to a scalar multiplier.

Eq. 6 defines a family of asymptotically unbiased estimators parameterized by the choice of functions $\alpha_{ij}(\mathbf{x})$, known in the statistics literature as *extended bridge sampling* estimators [19]. By making the choice

$$\alpha_{ij}(\mathbf{x}) = N_j \hat{c}_j^{-1} / \sum_{k=1}^K N_k \hat{c}_k^{-1} q_k(\mathbf{x}), \quad (7)$$

we obtain an estimator that has been proven to be optimal, in the sense that it has the lowest variance for a large class of choices of $\alpha_{ij}(\mathbf{x})$ [19]. This estimator is also asymptotically unbiased and guaranteed to have a unique solution (up to a multiplicative constant) [19], and can also be derived from maximum-likelihood measure theoretic methods [18] or cast as a reverse logistic regression problem [18, 23].

While a closed-form expression for the \hat{c}_i cannot be obtained from Eqs. 6 and 7, the \hat{c}_i can nevertheless be easily computed by any suitable method for solving systems of coupled nonlinear equations. A simple self-consistent iteration method and an efficient Newton-Raphson solver are described in Appendix C.

In the large sample limit, the error in the ratios \hat{c}_i/\hat{c}_j will be normally distributed [19], and the asymptotic covariance matrix $\Theta_{ij} = \text{cov}(\theta_i, \theta_j)$, where $\theta_i \equiv \ln \hat{c}_i$, can be estimated by [18]

$$\hat{\Theta} = \mathbf{W}^T (\mathbf{I}_N - \mathbf{W} \mathbf{N} \mathbf{W}^T)^+ \mathbf{W} \quad (8)$$

where \mathbf{I}_N is the $N \times N$ identity matrix (with $N = \sum_{i=1}^K N_i$ the total number of samples), and $\mathbf{N} = \text{diag}(N_1, N_2, \dots, N_K)$. The superscript $+$ denotes a suitable generalized inverse, such as the standard Moore-Penrose pseudoinverse, since the quantity in parentheses will be rank-deficient. \mathbf{W} denotes the $N \times K$ matrix of weights

$$W_{ni} = \hat{c}_i^{-1} \frac{q_i(\mathbf{x}_n)}{\sum_{k=1}^K N_k \hat{c}_k^{-1} q_k(\mathbf{x}_n)}. \quad (9)$$

The samples are now indexed by a single index $n = 1, \dots, N$, as the association of which samples \mathbf{x}_n came from which distribution $p_i(\mathbf{x})$ is no longer relevant. We note that this definition ensures $\sum_{n=1}^N W_{ni} = 1$ for all $i = 1, \dots, K$ and $\sum_{i=1}^K N_i W_{ni} = 1$ for all $n = 1, \dots, N$. The computational cost of evaluating the pseudoinverse of an $N \times N$ matrix in computing $\hat{\Theta}$ can be reduced to that of computing the eigenvalue decomposition of a $K \times K$ matrix, and in many cases can be reduced directly to a simpler $K \times K$ matrix (see Appendix D).

The covariance of estimates of arbitrary functions $\phi(\theta_1, \dots, \theta_K)$ and $\psi(\theta_1, \dots, \theta_K)$ of the log normalization constants θ_i can be estimated from $\hat{\Theta}$ by the expansion

$$\text{cov}(\hat{\phi}, \hat{\psi}) = \sum_{i,j=1}^K \frac{\partial \phi}{\partial \theta_i} \hat{\Theta}_{ij} \frac{\partial \psi}{\partial \theta_j}. \quad (10)$$

III. FREE ENERGIES

When configurations are sampled with Boltzmann statistics, where $q_i(\mathbf{x}) \equiv \exp[-u_i(\mathbf{x})]$, Eqs. 6 and 7 produce the following estimating equations for the dimensionless free energies

$$\hat{f}_i = -\ln \sum_{j=1}^K \sum_{n=1}^{N_j} \frac{\exp[-u_i(\mathbf{x}_{jn})]}{\sum_{k=1}^K N_k \exp[\hat{f}_k - u_k(\mathbf{x}_{jn})]} \quad (11)$$

which must be solved self-consistently for the \hat{f}_i . Again, because the normalization constants are only determined up to a multiplicative constant, the estimated free energies \hat{f}_i are determined uniquely only up to an additive constant, so only differences $\Delta \hat{f}_{ij} = \hat{f}_j - \hat{f}_i$ will be meaningful.

The uncertainty in the estimated free energy difference can be computed from Eqs. 8 and 10 as

$$\delta^2 \Delta \hat{f}_{ij} \equiv \text{cov}(-\ln \hat{c}_j / \hat{c}_i, -\ln \hat{c}_j / \hat{c}_i) = \hat{\Theta}_{ii} - 2\hat{\Theta}_{ij} + \hat{\Theta}_{jj}$$

Free energy differences and uncertainties between states not sampled are easily estimated by augmenting the set of states with additional reduced potentials $u_i(\mathbf{x})$ with the number of samples $N_i = 0$. For these unsampled states, no additional self-consistent estimation is required, so many such states can be estimated very efficiently.

IV. EQUILIBRIUM EXPECTATIONS

The equilibrium expectation of some mechanical observable $A(\mathbf{x})$ that depends only on configuration \mathbf{x} (and not momentum) is given by Eq. 4, and can be computed as a ratio of normalization constants c_A/c_a by defining two additional “states” characterized by the functions [22]

$$q_A(\mathbf{x}) = A(\mathbf{x}) q(\mathbf{x}) \quad ; \quad q_a(\mathbf{x}) = q(\mathbf{x})$$

where again $q(\mathbf{x}) \equiv \exp[-u(\mathbf{x})]$ if the expectation with respect to the Boltzmann weight is desired. Even though $q_A(\mathbf{x})$ may no longer be strictly nonnegative, we can still make use of the extended bridge sampling estimator (Eq. 6) to estimate the expectation $\langle A \rangle$ since $N_A = N_a = 0$.

Similarly, we augment the matrix \mathbf{W} (Eq. 9) with columns W_{nA} and W_{na} corresponding to $q_A(\mathbf{x})$ and $q_a(\mathbf{x})$, respectively:

$$\begin{aligned} W_{nA} &= \hat{c}_A^{-1} \frac{A(\mathbf{x}_n) \exp[-u(\mathbf{x}_n)]}{\sum_{k=1}^K N_k \exp[\hat{f}_k - u_k(\mathbf{x}_n)]} \\ W_{na} &= \hat{c}_a^{-1} \frac{\exp[-u(\mathbf{x}_n)]}{\sum_{k=1}^K N_k \exp[\hat{f}_k - u_k(\mathbf{x}_n)]} \end{aligned} \quad (12)$$

where the normalization constants \hat{c}_A and \hat{c}_a are defined in terms of self-consistent estimating equations as

$$\begin{aligned}\hat{c}_A &= \sum_{n=1}^N \frac{A(\mathbf{x}_n) \exp[-u(\mathbf{x}_n)]}{\sum_{k=1}^K N_k \exp[\hat{f}_k - u_k(\mathbf{x}_n)]} \\ \hat{c}_a &= \sum_{n=1}^N \frac{\exp[-u(\mathbf{x}_n)]}{\sum_{k=1}^K N_k \exp[\hat{f}_k - u_k(\mathbf{x}_n)]}\end{aligned}\quad (13)$$

We can then write the estimator of the expectation as

$$\hat{A} = \frac{\hat{c}_A}{\hat{c}_a} = \sum_{n=1}^N W_{na} A(\mathbf{x}_n)$$

and an estimator for the uncertainty as

$$\delta^2 \hat{A} \equiv \text{cov}(\hat{c}_A/\hat{c}_a, \hat{c}_A/\hat{c}_a) = \hat{A}^2(\hat{\Theta}_{AA} + \hat{\Theta}_{aa} - 2\hat{\Theta}_{Aa})$$

where the covariance matrix $\hat{\Theta}$ is now computed from the augmented \mathbf{W} . Covariances between estimates of $\langle A \rangle$ at different thermodynamic states, or between two observables $\langle A \rangle$ and $\langle B \rangle$, can also be constructed by adding the appropriate columns to the covariance matrix and applying Eq. 10 to estimate the desired uncertainty. If the dimensionless free energies \hat{f}_i have already been determined, computation of \hat{A} for any $A(\mathbf{x})$ and any $q(\mathbf{x})$ does not require additional iterative solution of the self-consistent estimating equations.

V. APPLICATION TO LABORATORY EXPERIMENTS

The MBAR estimator is not limited in application to data produced from simulation — it can also be applied to combine data from multiple equilibrium experiments in the presence of externally-applied fields. To illustrate, we estimate potential of mean force (PMF) of a DNA hairpin attached by dsDNA linkers to glass beads along the distance between the beads. The collection of equilibrium trajectories under a variety of constant force loads (corresponding to a linear external potential along the extension coordinate) for the DNA hairpin system 20R55/4T collected by an optical double trap experiment [24] we reported earlier [25]. The complete dataset was obtained from Michael Woodside (National Institute for Nanotechnology, NRC and Department of Physics, University of Alberta), and consists of 16 trajectories at 296.15 K, each 5 s in duration and sampled with a period of 0.1 ms, totaling 50 000 samples each. Each trajectory was collected under a different constant force load, with force loads ranging from 12.35 pN to 14.41 pN, with an estimated 10% relative error in the measurement of this force value.

The data were analyzed to produce an optimal estimate of the PMF under a force load of 14.19 pN, a

force load where it is difficult to determine the entire PMF to high precision from equilibrium trajectories at this force load alone (Figure 1). The sampled extension range was divided up into 50 unequally-sized bins such that the number of samples per bin was equal in order to avoid regions with zero histogram counts, as would occur with equally-spaced bins. Analysis with the MBAR estimator took 18 s on a standard 2.16 GHz Intel Core 2 Duo MacBook Pro, and the resulting error bars are more than an order of magnitude smaller than those derived from the single trajectory at this force load in the poorly sampled region of the PMF. Below, we describe how this analysis was performed with both types of analysis.

To estimate the potential of mean force from the 14.19 pN trajectory alone (black filled squares in Figure 1), the total number of counts N_i per histogram bin was determined, and the reduced potential of mean force (in units of kT) f_i computed up to an irrelevant additive constant from

$$f_i = -\ln(N_i/w_i) \quad (14)$$

where w_i is the relative width of bin i , necessary to correct for the nonuniform bin sizes. The statistical uncertainty in the histogram count was estimated by standard methods (see Eq. 26 of [26])

$$\begin{aligned}\delta^2 N_i &= g N_i (1 - N_i/N) \\ \delta^2 f_i &= \frac{\delta^2 N_i}{N_i^2}\end{aligned}\quad (15)$$

where g is the statistical inefficiency of the extension time series, estimated from the extension autocorrelation function (see Section 5.2 of [26]).

To estimate the potential of mean force using the MBAR estimator, the dataset was first subsampled with an interval equal to the statistical inefficiency of each trajectory at constant force to produce a set of uncorrelated samples. The reduced potential energy for each state k under the experimental conditions corresponds to

$$u_k(\mathbf{x}) = \beta[U_0(\mathbf{x}) + pV(\mathbf{x}) + U_k^{\text{ext}}(\mathbf{x})] \quad (16)$$

where $U_0(\mathbf{x})$ is the (unknown) potential energy function of the system in the absence of an externally-applied biasing potential, $pV(\mathbf{x})$ is the pressure-volume work, and $U_k^{\text{ext}}(\mathbf{x})$ is the (known) externally applied biasing potential, given by

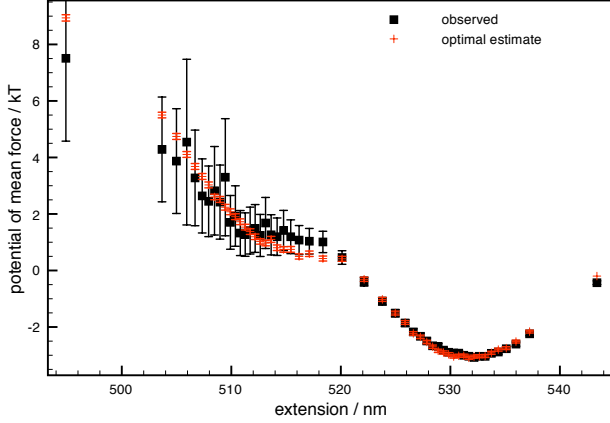
$$U_k^{\text{ext}}(\mathbf{x}) = -F_k z(\mathbf{x}) + a_k \quad (17)$$

where $z(\mathbf{x})$ is the extension coordinate, F_k is the constant applied force along the positive z direction, and a_k is a constant offset.

Because only differences $u_i(\mathbf{x}) - u_j(\mathbf{x})$ appear in the estimating equations (Eq. 11), the unknown components of the reduced potential energy cancel out and need not be considered

$$u_i(\mathbf{x}) - u_j(\mathbf{x}) = -\beta(F_i - F_j)z(\mathbf{x}) + \beta(a_i - a_j) \quad (18)$$

FIG. 1: Potential of mean force of DNA hairpin and dsDNA handles system 20R55/4T under 14.19 pN external force.



While the constant term $\beta(a_i - a_j)$ involving the unknown zero potential intercepts will appear in the estimated state free energies, these do not influence computed expectations in Eq. 14, and are hence irrelevant.

The probability of finding the system in a bin i under the conditions of interest is given by the expectation

$$p_i = \langle \chi_i(z) \rangle \quad (19)$$

where $\chi_i(z)$ is an indicator function that assumes the value of 1 if the system is in bin i and zero otherwise. The potential of mean force (in units of thermal energy kT) can then be computed from the p_i , up to an irrelevant additive constant, as

$$f_i = -\ln(p_i/w_i) \quad (20)$$

and the uncertainties propagated by Eq. 10.

Because each potential of mean force is only determined up to an arbitrary additive constant, the mean value of each PMF was subtracted before plotting. This is equivalent to choosing the additive constants so as to obtain an optimal least-squares RMS fit between the two PMFs.

It should be noted that these PMF corresponds to the potential of mean force along the entire system connected to the glass beads, which includes not only the DNA hairpin but the two dsDNA linkers and their attachments to the glass beads. In other work, deconvolution or related methods have been applied to correct for the stretching of the linkers to estimate PMF for the DNA hairpin alone [27].

VI. DISCUSSION

The MBAR estimator presented here provides a rapid and robust way to extract estimates of free energy differences and equilibrium expectations from multiple equilibrium samples of different thermodynamic states in a

statistically optimal way. As the estimator is asymptotically efficient among a wide class of “bridge sampling” estimators [19], which includes EXP and BAR as members, the resulting estimates from MBAR will have the lowest (or equal) variance in the large sample limit.

While multiple histogram techniques [10, 11, 12, 14] have been widely used for combining data from multi-state simulations, the MBAR estimator supplants these methods in the majority of cases. Most importantly, it provides a reliable and inexpensive method for estimating the uncertainties in the resulting estimates and their correlations, which are critical for propagating uncertainties to quantities of interest. Additionally, the elimination of histograms avoids both the bias arising from discretization of continuous energies as well as the computational overhead of constructing and storing high dimensional histograms.

In this framework, multiple histogram reweighting methods such as WHAM can be understood as a histogram kernel density estimator approximation to MBAR. In some applications, histograms can reduce the computational expense required for solving the estimating equations (Eq. 6) at the expense of introducing bias. When samples are distributed according to the Boltzmann weight, the estimator for the free energies (Eq. 11) is precisely Eq. 21 of [11] or Eq. 15 of [28], in both cases presented as a reduction of the histogram bin width to zero in the standard WHAM equations (Eqs. 19–20 of [11]). While the validity of this limit is dubious — the derivations in these references rely upon an estimate of the uncertainty in each histogram count which cannot be correct when the bins are nearly empty — the derivation of this equation from the extended bridge sampling estimator demonstrates for the first time that these equations are, in fact, asymptotically unbiased estimators of the true free energy differences.

The MBAR estimator also can be considered a multistate generalization of Bennett’s acceptance ratio estimator (BAR) [8]. In deriving BAR, Bennett constructed an estimator from Eq. 5 directly, determining the single $\alpha(x)$ which minimized the variance of the estimator of the free energy difference between only two states. In deriving MBAR, summing over all states j and determining the functions $\alpha_{ij}(x)$ that minimizes the covariance matrix of the estimator for ratios of normalization constants produces an optimal estimator for the multistate case. A proof of the equivalence of MBAR and BAR for two states can be found in Appendix E.

BAR and a recent pairwise multistate generalization (which we shall refer to as PBAR) [29] differ from MBAR in that they can also be applied to *nonequilibrium* work measurements between pairs of states, in addition to equilibrium reduced potential differences (*instantaneous* work measurements). However, PBAR constructs a total likelihood function from products of likelihood functions connecting pairs of states, assuming independence of all work measurements. For equilibrium samples, this means that a sampled configuration x_n from a state

i can only be used to provide information about the instantaneous work required to switch to a *single* other state j for use in the PBAR estimator, whereas in MBAR, each sampled x_n can be used to provide information about *all* states. As a result, MBAR should require significant fewer samples from each state to produce an estimate of equivalent precision with equilibrium data.

A Python implementation of the MBAR estimator described here is available under the GNU General Public License (GPL), and is provided online, along with several example applications, at <https://simtk.org/home/pymbar>.

VII. ACKNOWLEDGMENTS

We are indebted to Michael Woodside for providing us with detailed datasets from Ref. [25] and help-

ful comments. We thank Evangelos A. Coutsias, Gavin E. Crooks, Fernando A. Escobedo, Edward H. Feng, Andrew Gelman, Andrew I. Jewett, Libusha Kelly, Jun S. Liu, David D. L. Minh, David L. Mobley, Frank M. Noé, Vijay S. Pande, Sanghyun Park, M. Scott Shell, Zhiqiang Tan, Matthew A. Wyczalkowski, and the anonymous reviewers for enlightening discussions and constructive comments on this manuscript. JDC gratefully acknowledges support from Ken A. Dill through NIH grant GM34993 and Vijay S. Pande through an NSF grant for Cyberinfrastructure (NSF CHE-0535616), and MRS support from Richard A. Friesner and an NIH NRSA Fellowship.

APPENDIX A: CORRELATED TIME SERIES DATA

While estimating equations based on Eq. 6 can be applied to correlated or uncorrelated datasets, provided that the empirical estimator

$$\langle A \rangle_i \approx \frac{1}{N_i} \sum_{n=1}^{N_i} A(x_{in}) \quad (\text{A1})$$

remains asymptotically unbiased, the asymptotic covariance matrix estimator (Eq. 8) only produces sensible estimates when applied to *uncorrelated* datasets. Application to correlated datasets may produce severe underestimates of the true statistical uncertainty, and should be avoided.

A set of uncorrelated configurations can be obtained from a correlated time series, such as is generated by a molecular dynamics or Metropolis Monte Carlo simulation, by subsampling the time series with an interval approximately equal to the equilibrium relaxation time for the system. Because the equilibrium relaxation time is difficult to compute for all but the simplest systems, we find the *maximum* of the statistical inefficiency g computed for several relevant observables (such as the reduced potential $u_k(x)$ in Boltzmann-weighted sampling, structural observables $A(x)$ in the computation of potentials of mean force, etc.) provides a practical estimate useful for subsampling.

The statistical inefficiency g_A of the observable $A(x)$ of a time series $\{x_t\}_{t=1}^T$ is formally defined as (see Janke [30] for a detailed exposition)

$$\begin{aligned} g_A &\equiv 1 + 2\tau_A \\ \tau_A &\equiv \sum_{t=1}^T (1 - t/T) C_{AA}(t) \\ C_{AA}(t) &\equiv \frac{\langle A(x_{t_0}) A(x_{t_0+t}) \rangle_i - \langle A \rangle_i^2}{\langle A^2 \rangle_i - \langle A \rangle_i^2} \end{aligned} \quad (\text{A2})$$

where τ_A denotes the integrated autocorrelation time and $C_{AA}(t)$ the normalized fluctuation autocorrelation function of the observable A .

Direct application of these equations substituting the empirical estimator for the expectation can be problematic due to statistical noise. As a result, there exist a number of standard procedures [26, 30, 31, 32] to improve the quality and stability of this estimate for physical systems, making use of properties such as stationarity. The fast method for estimating the integrated autocorrelation time described in Section 5.2 of Chodera et al. [26] is implemented in the Python implementation of MBAR available online.

APPENDIX B: RECOMPUTATION OF REDUCED ENERGIES AT MULTIPLE STATES

Application of MBAR to simulation data requires $u_k(\mathbf{x}_n)$ to be evaluated for all K reduced potential functions $u_k(\mathbf{x})$ and all N uncorrelated sampled configuration \mathbf{x}_n , a total of KN reduced potential evaluations. In practice, this is not overly burdensome; the samples \mathbf{x}_n are generally produced by schemes that generate chains of highly correlated samples, such as molecular dynamics or Monte Carlo simulations. Once the stored configurations are sub-sampled to eliminate correlations and produce an effectively uncorrelated sample (as described above), the number of remaining samples $N \approx T/g$ is generally smaller than the number of samples T produced during the simulation by one or more orders of magnitude.

In cases where the $u_k(\mathbf{x})$ differ only by a linear scaling parameter of one or more components (such as temperature or an external field parameter), computation of $u_k(\mathbf{x}_n)$ for all K states is a trivial operation. For other cases, such as when all samples are collected from thermodynamic states that only differ in the external biasing potential (e.g. linear or harmonic), we note that the reduced energy differences $u_k(\mathbf{x}) - u_i(\mathbf{x})$ involve only differences in the external biasing potential, which can often be rapidly computed. Section V contains an illustration of this in application to single-molecule pulling experiments.

APPENDIX C: EFFICIENT SOLUTION OF THE ESTIMATING EQUATIONS

A number of methods can be used to obtain a self-consistent solution to the free energy estimating equations obtained from combining Eqs. 6 and 7

$$\hat{c}_i = \sum_{j=1}^K \sum_{n=1}^{N_j} \frac{q_i(\mathbf{x}_{jn})}{\sum_{k=1}^K N_k \hat{c}_k^{-1} q_k(\mathbf{x}_{jn})}. \quad (\text{C1})$$

or, in terms of the dimensionless free energies $f_i = -\ln c_i$,

$$\hat{f}_i = -\ln \sum_{j=1}^K \sum_{n=1}^{N_j} \frac{q_i(\mathbf{x}_{jn})}{\sum_{k=1}^K N_k e^{\hat{f}_k} q_k(\mathbf{x}_{jn})}. \quad (\text{C2})$$

While any method capable of solving a coupled set of nonlinear equations may be employed here, we describe some practical choices we made in the implementation of this algorithm. While any method capable of solving a coupled set of nonlinear equations may be employed, we describe two approaches to their solution: a straightforward yet reliable self-consistent iteration method and an efficient yet slightly less reliable Newton-Raphson method. Both methods are implemented in the Python implementation of the estimator available online.

1. Self-consistent iteration

As in [11], the \hat{f}_i could be obtained by self-consistent iteration of Eq. 11 using the last set of iterates $\{\hat{f}_i^{(n)}\}_{i=1}^K$ to produce a new estimated set of iterates $\{\hat{f}_i^{(n+1)}\}_{i=1}^K$:

$$\hat{f}_i^{(n+1)} = -\ln \sum_{j=1}^K \sum_{n=1}^{N_j} \frac{q_i(\mathbf{x}_{jn})}{\sum_{k=1}^K N_k e^{\hat{f}_k^{(n)}} q_k(\mathbf{x}_{jn})}. \quad (\text{C3})$$

Convergence is assured regardless of the initial choice of $f_i^{(0)}$, so it is sufficient to initialize the iteration by choosing $f_i^{(0)} = 0$. Alternative initial choices of the initial reduced free energies $f_i^{(0)}$ may speed convergence. For example, we have found the choice

$$f_k^{(0)} = \frac{1}{N_k} \sum_{n=1}^{N_k} \ln q(\mathbf{x}_{kn}) \quad (\text{C4})$$

which, for Boltzmann weighting ($q_k(\mathbf{x}) = \exp[-u_k(\mathbf{x})]$) corresponds to the average reduced potential energy, usually works well. Additional inexpensive choices are possible, such as fixing $f_1^{(0)} = 0$ and estimating consecutive differences $(f_{k+1}^{(0)} - f_k^{(0)})$, $k = 1, 2, \dots, K-1$ using the Bennett acceptance ratio (BAR) estimator [8, 9].

a. Cautions

For numerical reasons, it is convenient to constrain $f_1 = 0$ during the course of iteration by subtracting f_1 from the updated values in order to obtain a unique solution and prevent uncontrolled growth in the magnitude of the estimates. Iteration is terminated when the quantities of interest change by a fraction of the desired precision with additional iterations, but a convenient rule of thumb is to terminate when $\max_{i=2,\dots,K} |f_i^{(n+1)} - f_i^{(n)}| / |f_i^{(n)}| < 10^{-7}$. Because the quantities of interest and the relative free energies can converge at different rates, it is advised that the former be monitored when possible.

It is also critical to avoid *overflow* in the computation of exponentials e^a . To compute log sums of the form $\ln \sum_{n=1}^N e^{a_n}$, we can use the equivalent form

$$\ln \sum_{n=1}^N \exp[a_n] = c + \ln \sum_{n=1}^N \exp[a_n - c] \quad (\text{C5})$$

where $c \equiv \max_n a_n$. To minimize *underflow*, the terms $\exp[a_n - c]$ can be summed in order from smallest to largest.

2. Newton-Raphson

A more efficient approach to determination of the \hat{f}_i is to employ a Newton-Raphson solver, which has the advantage of quadratic convergence (a near doubling of the number of digits of precision) with each iteration when sufficiently near the solution. Because each iteration requires inversion of a $(K-1) \times (K-1)$ matrix, this approach is only efficient if K is small, say $K < 100$, but this will be satisfied in a wide number of cases.

First, we write the estimating equations in terms of a set of functions $g_i(\theta)$ such that the solution of the estimating equations (Eq. C2) corresponds to $\mathbf{g}(\hat{\theta}) = \mathbf{0}$. Several such choices of both the function $\mathbf{g}(\theta)$ and the parameterization (the normalization constants c_i or their logarithms θ_i) are possible, and the efficiencies of approaches based on different choices may differ substantially, but we find it convenient to choose

$$g_i(\theta) = N_i - N_i \sum_{n=1}^N W_{ni}(\theta). \quad (\text{C6})$$

where W_{ni} is defined in Eq. 9. It can easily be seen that $\mathbf{g}(\hat{\theta}) = \mathbf{0}$ is equivalent to the estimating equations:

$$\begin{aligned} g_i(\hat{\theta}) &= 0 \\ \Leftrightarrow N_i - N_i \sum_{n=1}^N W_{ni}(\hat{\theta}) &= 0 \\ \Leftrightarrow N_i - N_i \frac{\hat{c}_i^{-1} q_i(\mathbf{x}_n)}{\sum_{k=1}^K N_k \hat{c}_k^{-1} q_k(\mathbf{x}_n)} &= 0 \\ \Leftrightarrow \hat{c}_i &= \frac{q_i(\mathbf{x}_n)}{\sum_{k=1}^K N_k \hat{c}_k^{-1} q_k(\mathbf{x}_n)} \end{aligned} \quad (\text{C7})$$

In Newton-Raphson, the function $\mathbf{g}(\theta)$ is expanded about the current iterate $\theta^{(n)}$ to first order:

$$\mathbf{g}(\theta^{(n+1)}) \approx \mathbf{g}(\theta^{(n)}) + \mathbf{H}(\theta^{(n)})(\theta^{(n+1)} - \theta^{(n)}). \quad (\text{C8})$$

where

$$H_{ij}(\theta) = \frac{\partial g_i(\theta)}{\partial \theta_j} = \begin{cases} - \sum_{n=1}^N N_i W_{ni}(1 - N_i W_{ni}) & \text{if } i = j \\ \sum_{n=1}^N N_i W_{ni} N_j W_{nj} & \text{if } i \neq j \end{cases} \quad (\text{C9})$$

We seek the next iterate $\theta^{(n+1)}$ such that $\mathbf{g}(\theta^{(n+1)}) = \mathbf{0}$, which yields the update equation

$$\theta^{(n+1)} = \gamma [\mathbf{H}(\theta^{(n)})]^+ \mathbf{g}(\theta^{(n)}) \quad (\text{C10})$$

where $^+$ denotes the pseudoinverse. If all the $q_i(\mathbf{x})$ are unique and $N_i > 0$ for all states, the standard matrix inverse may be substituted for the pseudoinverse.

a. Cautions

We only need to iterate over states for which $N_i > 0$; the relative free energies of states where $N_i = 0$, and expectation values at all states, can be determined after the self-consistent equations are solved to determine the relative free energies of states where $N_i > 0$. Since we must constrain $f_1 = 0$ to avoid drift during the process of free energy determination, we can simply use a modified form of Eq. C10 where rows and columns corresponding to the first state are omitted.

$$\theta_{(2:K, 2:K)}^{(n+1)} = \gamma [\mathbf{H}(\theta^{(n)})_{(2:K, 2:K)}]^+ \mathbf{g}(\theta^{(n)})_{(2:K, 2:K)}. \quad (\text{C11})$$

$\gamma \in (0, 1]$ is a scalar multiplier that controls the rate of convergence. Since the initial iterate $\theta^{(1)}$ may be far from the realm of quadratic convergence (i.e. outside the range at which the Taylor expansion in Eq. C8 holds), it is often safer to choose an initial $\gamma \ll 1$. We have found $\gamma = 0.1$ works well for the first step, with $\gamma = 1$ used thereafter.

Even then, there are times when with reduced γ does not prevent numerical instability. The instability may be due to the initial guess iterate $\theta^{(0)}$ being too far from the region of quadratic convergence, such that the first-order Taylor expansion above is a poor approximation to $\mathbf{g}(\theta)$ in Eq. C8. In this case, a better procedure for choosing the initial iterate may aid convergence. Starting with one or more iterations of the self-consistent method (Section C1), or using an initial estimate from application of BAR [9] to sequential states may be sufficient. Less commonly, failure to converge may be result from numerical precision limiting the accurate calculation of the pseudoinverse $[\mathbf{H}(\theta^{(n)})_{(2:K, 2:K)}]^+$. In all cases, we find that self-consistent iteration still works reliably to recover the estimator, and can be used as a fallback procedure.

APPENDIX D: EFFICIENT COMPUTATION OF THE ASYMPTOTIC COVARIANCE MATRIX

1. Singular value decomposition

The $N \times K$ matrix \mathbf{W} (Eq. 9 in the main paper) can be written in terms of its singular value decomposition

$$\mathbf{W} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \quad (\text{D1})$$

where \mathbf{U} is an $N \times N$ unitary matrix of left singular vectors (such that $\mathbf{U} \mathbf{U}^T = \mathbf{I}_N$), $\mathbf{\Sigma}$ is an $N \times K$ matrix containing $L < K$ singular values along the diagonal, and \mathbf{V} is a $K \times K$ unitary matrix of right singular vectors.

The estimator for the asymptotic covariance matrix $\hat{\Theta}$ (Eq. 8) can then be expanded to

$$\begin{aligned} \hat{\Theta} &= \mathbf{W}^T (\mathbf{I}_N - \mathbf{W} \mathbf{N} \mathbf{W}^T)^+ \mathbf{W} \\ &= (\mathbf{U} \mathbf{\Sigma} \mathbf{V}^T)^T [\mathbf{I}_N - (\mathbf{U} \mathbf{\Sigma} \mathbf{V}^T) \mathbf{N} (\mathbf{U} \mathbf{\Sigma} \mathbf{V}^T)^T]^+ (\mathbf{U} \mathbf{\Sigma} \mathbf{V}^T) \\ &= \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T [\mathbf{I}_N - \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \mathbf{N} \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T]^+ \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \\ &= \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T [\mathbf{U} \mathbf{U}^T - \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \mathbf{N} \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T]^+ \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \\ &= \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T [\mathbf{U} (\mathbf{I}_N - \mathbf{\Sigma} \mathbf{V}^T \mathbf{N} \mathbf{V} \mathbf{\Sigma}^T) \mathbf{U}^T]^+ \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \\ &= \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T \mathbf{U} [\mathbf{I}_N - \mathbf{\Sigma} \mathbf{V}^T \mathbf{N} \mathbf{V} \mathbf{\Sigma}^T]^+ \mathbf{U}^T \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \\ &= \mathbf{V} \mathbf{\Sigma}^T [\mathbf{I}_N - \mathbf{\Sigma} \mathbf{V}^T \mathbf{N} \mathbf{V} \mathbf{\Sigma}^T]^+ \mathbf{\Sigma} \mathbf{V}^T \end{aligned} \quad (\text{D2})$$

We partition the matrix of singular values $\mathbf{\Sigma}$ into a $K \times K$ diagonal region Σ_K (of which only the first $L \leq K$ diagonal entries will be nonzero) and an $(N - K) \times K$ zero matrix $\mathbf{0}$:

$$\mathbf{\Sigma} = \begin{bmatrix} \Sigma_K \\ \mathbf{0} \end{bmatrix} \quad (\text{D3})$$

We can then rewrite the above expression as

$$\begin{aligned}\hat{\Theta} &= \mathbf{V} \begin{bmatrix} \Sigma_K & \mathbf{0} \end{bmatrix} \left\{ \begin{bmatrix} \mathbf{I}_K & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{(N-K)} \end{bmatrix} - \begin{bmatrix} \Sigma_K \\ \mathbf{0} \end{bmatrix} \mathbf{V}^T \mathbf{N} \mathbf{V} \begin{bmatrix} \Sigma_K & \mathbf{0} \end{bmatrix} \right\}^+ \begin{bmatrix} \Sigma_K \\ \mathbf{0} \end{bmatrix} \mathbf{V}^T \\ &= \mathbf{V} \Sigma_K [\mathbf{I}_K - \Sigma_K \mathbf{V}^T \mathbf{N} \mathbf{V} \Sigma_K]^+ \Sigma_K \mathbf{V}^T\end{aligned}\quad (\text{D4})$$

We note that pseudoinversion of the quantity in brackets now only requires $\mathcal{O}(K^3)$ work, though this can be further reduced to $\mathcal{O}(L^3)$ work if the reduced SVD is used.

The singular values Σ_K and matrix of right singular vectors \mathbf{V} can easily be computed from the eigenvalue decomposition of $\mathbf{W}^T \mathbf{W}$:

$$\begin{aligned}\mathbf{W}^T \mathbf{W} &= (\mathbf{U} \Sigma \mathbf{V}^T)^T (\mathbf{U} \Sigma \mathbf{V}^T) \\ &= \mathbf{V} \Sigma^T \mathbf{U}^T \mathbf{U} \Sigma \mathbf{V}^T \\ &= \mathbf{V} (\Sigma^T \Sigma) \mathbf{V}^T\end{aligned}\quad (\text{D5})$$

2. When \mathbf{W} has full column rank

In the case that \mathbf{W} has full column rank (because all $q_k(x)$, $k = 1, \dots, K$ are unique) we can make further progress. Using Eq. D4, we can write

$$\begin{aligned}\hat{\Theta} &= [\mathbf{V} \Sigma_K^{-1} (\mathbf{I}_K - \Sigma_K \mathbf{V}^T \mathbf{N} \mathbf{V} \Sigma_K + \mathbf{1} \mathbf{1}^T) \Sigma_K^{-1} \mathbf{V}^T]^{-1} \\ &= [\mathbf{V} (\Sigma_K^{-2}) \mathbf{V}^T - \mathbf{N}]^+ \\ &= [(\mathbf{W}^T \mathbf{W})^{-1} - \mathbf{N}]^+\end{aligned}\quad (\text{D6})$$

We note that $\mathbf{W}^T \mathbf{1}_N = \mathbf{1}_K$, and $\mathbf{W} \mathbf{N} \mathbf{1}_K = \mathbf{1}_N$, and so $\mathbf{W}^T \mathbf{W} \mathbf{N} \mathbf{1}_K = \mathbf{1}_K$, and observe that $[(\mathbf{W}^T \mathbf{W})^{-1} - \mathbf{N}]$ has rank $K - 1$ with kernel $\mathbf{1}_K$

$$\begin{aligned}[(\mathbf{W}^T \mathbf{W})^{-1} - \mathbf{N}] \mathbf{1}_K &= (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{1}_K - \mathbf{N} \mathbf{1}_K \\ &= (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{W} \mathbf{N} \mathbf{1}_K - \mathbf{N} \mathbf{1}_K \\ &= \mathbf{N} \mathbf{1}_K - \mathbf{N} \mathbf{1}_K \\ &= \mathbf{0}\end{aligned}$$

We can supplement the quantity in brackets with $b \mathbf{1}_K \mathbf{1}_K^T$, where b is some nonzero scalar, without changing the covariance values computed from it, and make it invertible:

$$\hat{\Theta} = [(\mathbf{W}^T \mathbf{W})^{-1} - \mathbf{N} + b \mathbf{1}_K \mathbf{1}_K^T]^{-1} \quad (\text{D7})$$

We choose $b = N^{-1}$ to ensure the inversion is well-conditioned (as in [18]), producing

$$\hat{\Theta} = [(\mathbf{W}^T \mathbf{W})^{-1} - \mathbf{N} + \mathbf{1}_K \mathbf{1}_K^T / N]^{-1}. \quad (\text{D8})$$

APPENDIX E: EQUIVALENCE OF MBAR AND BAR FOR TWO STATES

We start with Eq. 11. For ease of use, we define $\Delta\hat{f} = \hat{f}_2 - \hat{f}_1$ and $\Delta u(\mathbf{x}) = u_2(\mathbf{x}) - u_1(\mathbf{x})$ and $M = \ln N_2/N_1$. Without loss of generalization, since the equations are symmetric, we examine the self-consistent equation for \hat{f}_1 .

$$\begin{aligned}
 \hat{f}_1 &= -\ln \sum_{j=1}^2 \sum_{n=1}^{N_j} \frac{\exp[-u_1(\mathbf{x}_{jn})]}{\sum_{k=1}^2 N_k \exp[\hat{f}_k - u_k(\mathbf{x}_{jn})]} \\
 1 &= \sum_{n=1}^{N_1} \frac{\exp[\hat{f}_1 - u_1(\mathbf{x}_{1n})]}{N_1 \exp[\hat{f}_1 - u_1(\mathbf{x}_{1n})] + N_2 \exp[\hat{f}_2 - u_2(\mathbf{x}_{1n})]} + \\
 &\quad \sum_{n=1}^{N_2} \frac{\exp[\hat{f}_1 - u_1(\mathbf{x}_{2n})]}{N_1 \exp[\hat{f}_1 - u_1(\mathbf{x}_{2n})] + N_2 \exp[\hat{f}_2 - u_2(\mathbf{x}_{2n})]} \\
 1 &= \sum_{n=1}^{N_1} \frac{1}{N_1 + N_2 \exp[\Delta\hat{f} - \Delta u(\mathbf{x}_{1n})]} + \sum_{n=1}^{N_2} \frac{1}{N_1 + N_2 \exp[\Delta\hat{f} - \Delta u(\mathbf{x}_{2n})]} \\
 N_1 &= \sum_{n=1}^{N_1} \frac{1}{1 + \frac{N_2}{N_1} \exp[\Delta\hat{f} - \Delta u(\mathbf{x}_{1n})]} + \sum_{n=1}^{N_2} \frac{1}{1 + \frac{N_2}{N_1} \exp[\Delta\hat{f} - \Delta u(\mathbf{x}_{2n})]} \\
 N_1 &= \sum_{n=1}^{N_1} \frac{1}{1 + \exp[M + \Delta\hat{f} - \Delta u(\mathbf{x}_{1n})]} + \sum_{n=1}^{N_2} \frac{1}{1 + \exp[M + \Delta\hat{f} - \Delta u(\mathbf{x}_{2n})]} \tag{E1}
 \end{aligned}$$

We make the additional observation that

$$\frac{1}{1 + \exp(x)} - 1 = -\frac{1}{1 + \exp(-x)}$$

which allows us to write Eq. E1 as:

$$\begin{aligned}
 0 &= \sum_{n=1}^{N_1} \left[\frac{1}{1 + \exp[M + \Delta\hat{f} - \Delta u(\mathbf{x}_{1n})]} - 1 \right] + \sum_{n=1}^{N_2} \frac{1}{1 + \exp[M + \Delta\hat{f} - \Delta u(\mathbf{x}_{2n})]} \\
 &= \sum_{n=1}^{N_1} \frac{1}{1 + \exp[M + \Delta\hat{f} - \Delta u(\mathbf{x}_{1n})]} - \sum_{n=1}^{N_2} \frac{1}{1 + \exp[-M - \Delta\hat{f} + \Delta u(\mathbf{x}_{2n})]}
 \end{aligned}$$

Which is exactly the equation for BAR presented in Shirts *et al.* [9].

We now examine the expression for the variance limited to two states. When the two thermodynamic states are not identical, the \mathbf{W} will have full rank, and the asymptotic covariance matrix can be written as (see Eq. D7 above):

$$\Theta = [(\mathbf{W}^T \mathbf{W})^{-1} - \mathbf{N} + \mathbf{1}_K \mathbf{1}_K^T / N]^{-1}$$

where we have from Eq. 9

$$W_{ni} = \frac{\exp(\hat{f}_i - u_i(\mathbf{x}_n))}{\sum_{k=1}^K N_k \exp(\hat{f}_k - u_k(\mathbf{x}_n))}$$

Defining F as the Fermi function, and $X_n = M + \Delta\hat{f} - \Delta u(\mathbf{x}_n)$, then in the case of two states $W_{n1} = N_1^{-1} F(X_n)$ and $W_{n2} = N_2^{-1} F(-X_n)$.

The matrix $\mathbf{W}^T \mathbf{W}$ can then be written as:

$$\mathbf{W}^T \mathbf{W} = \sum_{n=1}^N \begin{pmatrix} N_1^{-2} F(X_n)^2 & N_1^{-1} N_2^{-1} F(X_n) F(-X_n) \\ N_1^{-1} N_2^{-1} F(X_n) F(-X_n) & N_2^{-2} F(-X_n)^2 \end{pmatrix} \tag{E2}$$

If we represent the matrix $(\mathbf{W}^T \mathbf{W})_{ij} = a_{ij}$, the determinant $|\mathbf{W}^T \mathbf{W}|$ will be $D = a_{11}a_{22} - a_{21}a_{12}$. The variance of ratios is actually independent of multiplicative factor used in front of $\mathbf{1}_K \mathbf{1}_K$, as we will show below, so we will use b in place of $1/N$ for generality. The inverse of the covariance matrix is then:

$$\Theta^{-1} = (\mathbf{W}^T \mathbf{W})^{-1} - \mathbf{N} + b \mathbf{1}_N \mathbf{1}_N^T = \begin{pmatrix} \frac{a_{22}}{D} - N_1 + b & -\frac{a_{21}}{D} + b \\ -\frac{a_{12}}{D} + b & \frac{a_{11}}{D} - N_2 + b \end{pmatrix}$$

The determinant will then be:

$$|\Theta^{-1}| = \frac{1}{D} - \frac{N_2 a_{22} - N_1 a_{11}}{D} + N_1 N_2 - b(N_1 + N_2) + b \frac{a_{11} + a_{22} + a_{12} + a_{21}}{D} \quad (\text{E3})$$

However, we note that $(\mathbf{W}^T \mathbf{W})^{-1} - \mathbf{N}$ is singular, and thus the sum of the first three terms in Eq. E3 equals zero. Additionally, because it has kernel $\mathbf{1}_K$, it must also satisfy $a_{22} - a_{21} - N_1 D = 0$ and $a_{11} - a_{12} - N_2 D = 0$. Because we know by symmetry that $a_{12} = a_{21}$, which we denote by simply a , this determinant then becomes:

$$|\Theta^{-1}| = \frac{b}{D} [a_{11} + a_{22} + a_{12} + a_{21} - (N_1 + N_2)D] = \frac{4ab}{D}$$

We then obtain:

$$\Theta = [(\mathbf{W}^T \mathbf{W})^{-1} - \mathbf{N} + b \mathbf{1}_N \mathbf{1}_N^T]^{-1} = \frac{D}{4ab} \begin{pmatrix} \frac{a_{11}}{D} - N_2 + b & \frac{a}{D} - b \\ \frac{a}{D} - b & \frac{a_{22}}{D} - N_1 + b \end{pmatrix}$$

The variance in $f_1 - f_2$ will be $\Theta_{11} + \Theta_{22} - 2\Theta_{12}$, which reduces to:

$$\begin{aligned} \text{Var}(f_1 - f_2) &= \frac{D}{4ab} \left[\frac{a_{11} - N_1 D + bD + a_{22} - N_2 D + bD - 2a + 2bD}{D} \right] \\ &= \frac{1}{4ba} [(a_{11} - a - N_2 D) + (a_{22} - a - N_1 D) + 4bD] \\ &= \frac{D}{a} \end{aligned}$$

Which is indeed independent of $b \neq 0$. Since $a_{22} = N_1 + a$ and $a_{11} = N_2 + a$, given $D = a_{11}a_{22} - a^2$ (as noted above), we can find that $D = N_1^{-1}N_2^{-1}(1 - Na)$. We then obtain:

$$\begin{aligned} \text{Var}(\hat{f}_1 - \hat{f}_2) &= \frac{a_{11}a_{22} - a^2}{a} \\ &= \frac{1 - Na}{N_1 N_2 a} \\ &= \frac{1}{N_1 N_2 a} - \frac{N}{N_1 N_2} \\ &= \frac{1}{\sum_{i=1}^N F(X_n) F(-X_n)} - \frac{N}{N_1 N_2} \\ &= \left[\sum_{i=1}^N \frac{1}{2 + 2 \cosh(X_n)} \right]^{-1} - \frac{N}{N_1 N_2} \\ &= \frac{1}{N} \left[\left\langle \frac{1}{2 + 2 \cosh(M + \Delta \hat{f} - \Delta u(\mathbf{x}))} \right\rangle^{-1} - \left(\frac{N}{N_2} + \frac{N}{N_1} \right) \right] \end{aligned}$$

This is the equation for the asymptotic covariance of of free energies given for the BAR method in Shirts *et al.* [9]

[1] Here, a *thermodynamic state* is defined by a combination of potential energy function (including any bias-

ing potentials) and external thermodynamic parameters,

such as temperature, pressure, and chemical potential, all within the same thermodynamic ensemble (e.g. canonical, isothermal-isobaric, (semi)grand canonical).

- [2] G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **23**, 187 (1977).
- [3] E. Marinari and G. Parisi, *Europhys. Lett.* **19**, 451 (1992).
- [4] K. Hukushimi and K. Nemoto, *J. Phys. Soc. Jpn.* **65**, 1604 (1996).
- [5] R. W. Zwanzig, *J. Chem. Phys.* **22**, 1420 (1954).
- [6] B. Widom, *J. Chem. Phys.* **39**, 2808 (1963).
- [7] M. R. Shirts and V. S. Pande, *J. Chem. Phys.* **122**, 144107 (2005).
- [8] C. H. Bennett, *J. Comp. Phys.* **22**, 245 (1976).
- [9] M. R. Shirts, E. Bair, G. Hooker, and V. S. Pande, *Phys. Rev. Lett.* **91**, 140601 (2003).
- [10] A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.* **63**, 1195 (1989).
- [11] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, *J. Comput. Chem.* **13**, 1011 (1992).
- [12] C. Bartels and M. Karplus, *J. Comput. Chem.* **18**, 1450 (1997).
- [13] M. N. Kobrak, *J. Comput. Chem.* **24**, 1437 (2003).
- [14] E. Gallicchio, M. Andrec, A. K. Felts, and R. M. Levy, *J. Phys. Chem. B* **109**, 6722 (2005).
- [15] S. Park, D. L. Ensign, and V. S. Pande, *Phys. Rev. E* **74**, 066703 (2006).
- [16] Y. Vardi, *Ann. Stat.* **13**, 178 (1985).
- [17] R. D. Gill, Y. Vardi, and J. A. Wellner, *Ann. Stat.* **16**, 1069 (1988).
- [18] A. Kong, P. McCullagh, X.-L. Meng, D. Nicolae, and Z. Tan, *J. Royal Stat. Soc. B.* **65**, 585 (2003).
- [19] Z. Tan, *J. Am. Stat. Assoc.* **99**, 1027 (2004).
- [20] E. Mezei, *J. Comp. Phys.* **68**, 237 (1987).
- [21] C. Tsallis, *J. Stat. Phys.* **52**, 479 (1988).
- [22] Honi Doss makes this suggestion in the conference discussion of [18].
- [23] C. J. Geyer, Technical Report No. 568, School of Statistics, University of Minnesota, Minneapolis, Minnesota (unpublished).
- [24] W. J. Greenleaf, M. T. Woodside, E. A. Abbondanzieri, and S. M. Block, *Phys. Rev. Lett.* **95**, 208102 (2005).
- [25] M. T. Woodside, W. M. Behnke-Parks, K. Larizadeh, K. Travers, D. Herschlag, and S. M. Block, *Proc. Nat. Acad. Sci. USA* **103**, 6190 (2006).
- [26] J. D. Chodera, W. C. Swope, J. W. Pitera, C. Seok, and K. A. Dill, *J. Chem. Theor. Comput.* **3**, 26 (2007).
- [27] M. T. Woodside, P. C. Anthony, W. M. Behnke-Parks, K. Larizadeh, D. Herschlag, and S. M. Block, *Science* **314**, 1001 (2006).
- [28] M. Souaille and B. Roux, *Comp. Phys. Commun.* **135**, 40 (2001).
- [29] P. Maragakis, M. Spichty, and M. Karplus, *Phys. Rev. Lett.* **96**, 100602 (2006).
- [30] W. Janke, in *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, edited by J. Groten-dorst, D. Marx, and A. Murmatsu (John von Neumann Institute for Computing, Jülich, Germany, 2002), Vol. 10, pp. 423–445.
- [31] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, *J. Chem. Phys.* **76**, 637 (1982).
- [32] H. Flyvbjerg and H. G. Petersen, *J. Chem. Phys.* **91**, 461 (1989).