PHYSICAL REVIEW LETTERS

Optimal Estimates of Free Energies from Multistate Nonequilibrium Work Data

Paul Maragakis, Martin Spichty, and Martin Karplus

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA,
and Laboratoire de Chimie Biophysique, Institut de Science et d'Ingénierie Supramoléculaires, Université Louis Pasteur,
BP 70028, 8 allée Gaspard Monge, F-67083 Strasbourg Cedex, France

We derive the optimal estimates of the free energies of an arbitrary number of thermodynamic states from nonequilibrium work measurements; the work data are collected from forward and reverse switching processes and obey a fluctuation theorem. The maximum likelihood formulation properly reweights all pathways contributing to a free energy difference and is directly applicable to simulations and experiments. We demonstrate dramatic gains in efficiency by combining the analysis with parallel tempering simulations for alchemical mutations of model amino acids.

Introduction.—Free energy simulations play an ever increasing role in studies of condensed matter, particularly those concerned with problems in molecular biophysics. With the advent of accurate force fields, increases in computer power, and continuous methodological advances, "free energy simulations [have] come of age [1]." Because of the flexibility of the design and control of the simulations at the atomistic level, they can provide information not available from experiment. In parallel developments, a new generation of single molecule studies [2,3] are using advances in theory [4–8] to extract free energy information from the experiments. A recent example is a single molecule pulling experiment [2] that determined the folding free energies of RNA strands. The measured unfolding and refolding nonequilibrium work data [6] were analyzed with a generalization of Bennett's acceptance ratio method [4] to finite switching [7]. The Bennett method and its generalization were shown recently to provide the maximum likelihood estimator of the free energy difference, given a set of work data between two states [9]. Surprisingly, the Bennett method, which dates from 1976, has rarely been used in computations of free energy differences in biological systems; calculations have been based primarily on the exponential difference formula of Zwanzig [10] or the thermodynamic integration approach of Kirkwood [11].

Maximum likelihood estimators, under very general and verifiable conditions, are asymptotically consistent and efficient estimators [12]; i.e., they provide the smallest variance of any unbiased estimate of the parameters underlying the distribution of a large set of sampled data. In this Letter, we provide the maximum likelihood estimator of free energy differences from nonequilibrium work data sets for *multiple* states and verify that it is asymptotically unbiased and efficient. We also show that the method can be used to combine all of the sampled data from parallel tempering simulations [13,14] in an optimum way to obtain free energy differences. The analysis is directly applicable to present-day simulations of free energy differences

of interest in molecular biophysics, e.g., the properties of mutants or the study of multiligand binding. In what follows, we first derive the method and then demonstrate its utility in the analysis of parallel tempering simulations of alchemical mutations of model amino acids.

Derivation.—Consider a set of N thermodynamic states, which correspond to systems with different Hamiltonian, possibly sampled at different external conditions (e.g., variation of the temperature). A switch from one state to another can be produced by either an instantaneous (sudden) change of the Hamiltonian and the external conditions with the two systems in the same microstate [4,10] or a sequence of gradual changes that lead from the Hamiltonian and external conditions of the initial state to the final state [5,6,11].

Consider the pair of thermodynamic states i and j and a forward and reverse switch, such that a fluctuation theorem [4,6,15] of the following form holds for the probability distributions of a path-dependent quantity W_{ij} that is odd under path reversal ($W_{ji} = -W_{ij}$):

$$p(W_{ij}|i \to j)e^{-W_{ij}} = p(W_{ij}|j \to i)e^{-A_{ij}}, \qquad (1)$$

with $p(W_{ij}|i \to j)$ the probability of measuring W_{ij} along a path sampled in the switch from i to j, and A_{ij} an appropriate state-dependent quantity; W_{ij} can be thought of as a generalization of the work and A_{ij} as the free energy difference. In the instantaneous switch case, W_{ij} and A_{ij} are

$$W_{ij}(x) = \beta_j E_j(x) - \beta_i E_i(x), \qquad A_{ij} = \ln Z_i - \ln Z_j, \qquad (2)$$

with x a microstate (coordinates and momenta) sampled at equilibrium in the initial ensemble, E_j the energy of state j, β_j the inverse temperature of bath j, $\ln Z_j$ the logarithm of the partition function of state j; with these definitions Eq. (1) holds, as shown in Ref. [4]. In the gradual switch case performed in contact with a constant temperature heat bath, $W_{ij}(x) = \beta W$, with W the work along the pathway,

and $A_{ij} = \beta \Delta F_{ij}$, with ΔF_{ij} the Helmholtz free energy difference between states i and j. More general situations for which Eq. (1) holds are discussed in Refs. [6,15].

We follow the analysis of Shirts *et al.* [9] (see also Ref. [16]) to obtain the conditional probability that a work value W_{ij} along a path between states i and j resulted from a sampling of a forward $(i \rightarrow j)$ switching process. From the Bayes theorem [17], the ratio of probabilities of the forward to the backward directions of the switch given the work value and the end states i, j is

$$\frac{p(i \rightarrow j | W_{ij})}{p(j \rightarrow i | W_{ij})} = \frac{p(W_{ij} | i \rightarrow j) p(i \rightarrow j)}{p(W_{ij} | j \rightarrow i) p(j \rightarrow i)}, \qquad (3)$$

with $p(i \rightarrow j)$ the probability that the path between states i and j was sampled in the direction from i to j; for notational simplicity we omit the explicit dependence of all probabilities on the given pair of thermodynamic states $\{i, j\}$ and on A_{ij}. A given work value between the states i and j can be sampled in either the direction $i \rightarrow j$ or the direction $j \rightarrow i$, so that the numerator and denominator on the left-hand side of Eq. (3) sum to 1. Following Ref. [9], Eqs. (3) and (1) can be used to obtain

$$p(i \rightarrow j | W_{ij}) = f(-W_{ij} + A_{ij} + M_{ij}), \qquad (4)$$

$$M_{ij} = \ln n_{ij}^{tot} - \ln n_{ji}^{tot}, \qquad (5)$$

with f the Fermi function $f(x) = 1/(1 + e^x)$, and n_{ij}^{tot} the total number of uncorrelated work data in the direction $i \rightarrow j$. For the purpose of the maximum likelihood estimate of $\ln Z_i$, the ratio $p(i \rightarrow j)/p(j \rightarrow i)$ can be substituted with $n_{ij}^{tot}/n_{ji}^{tot}$ without loss of rigor [9,16]. Equation (4) resembles Bennett's acceptance ratio [4] of a switch move in a simultaneous Monte Carlo sampling of i and j that minimizes the variance of their free energy difference.

We can now write the joint likelihood p of observing forward switches from all states i to every other state j given the work data between these states as

$$p = \prod_i \prod_{j \neq i} \prod_{n_{ij}} f(-W_{ij,n_{ij}} + A_{ij} + M_{ij}), \qquad (6)$$

with $W_{ij,n_{ij}}$ the work values sampled along the n_{ij} paths for the switches $i \rightarrow j$. This equation, which follows from Eq. (4) for independent work data, also gives the probabilities of partitioning the work data for all pairs of states into those resulting from forward and from reverse switches, since the reverse process with $i > j$ corresponds to a forward process with $j > i$.

We now determine the set of $\ln Z_i$ that maximizes the probability in Eq. (6), or equivalently the logarithm of this probability. This is the central result of the present development. As can be seen from Eq. (2), Eq. (6) remains invariant when we multiply every partition function by the same constant. We can thus fix one of the $\ln Z_i$ (e.g., set $\ln Z_1$ to zero) and maximize the logarithm of Eq. (6) with respect to

the remaining ones. Since the derivatives of the objective function are available in closed form to all orders, we can use the Newton-Raphson method to search for a stationary point. Using the properties of the Fermi function, namely, $\partial \ln f(x)/\partial x = -f(-x)$ and $f'(x) = -1/(2 + 2 \cosh x)$, we obtain the first and second derivatives of $\ln p$:

$$F_a = \frac{\partial \ln p}{\partial \ln Z_a} = \sum_{ij} q_{ij}^a s_{ij},$$

$$s_{ij} = \sum_{n_{ij}} f(W_{ij,n_{ij}} - A_{ij} - M_{ij}), \qquad (7)$$

$$q_{ij}^a = -\frac{\partial A_{ij}}{\partial \ln Z_a} = \delta_{j,a} - \delta_{i,a},$$

$$H_{ab} = \frac{\partial^2 \ln p}{\partial \ln Z_a \partial \ln Z_b} = \sum_{ij} q_{ij}^a q_{ij}^b t_{ij},$$

$$t_{ij} = \sum_{n_{ij}} \frac{-1}{2 + 2 \cosh(W_{ij,n_{ij}} - A_{ij} - M_{ij})}, \qquad (8)$$

with a, b the indexes of the states with respect to which we obtain the derivatives and $\delta_{i,a}$ the Kronecker delta symbol. The full set of components of the forces F_a [Eq. (7)] and of the Hessian H_{ab} [Eq. (8)] can be efficiently calculated in a single scan through all pairs of states due to the sparseness of the array $\{q_{ij}^a\}$. Furthermore, the Hessian of Eq. (8) is always negative definite; i.e., direct substitution of q_{ij}^a, q_{ij}^b in Eq. (8) gives that for an arbitrary nonzero vector y of $\mathcal{R}^N$ the product $y^T H y = \sum_{ij} t_{ij} (y_i - y_j)^2$. This result is strictly smaller than zero if at least one of the $(y_i - y_j)$ is different from zero. The latter is always the case if we fix one of the components of y to zero, which is possible given the invariance of Eq. (6) to rescaling of all the partition functions. This means that the logarithm of Eq. (6) has only a single stationary point, which is a maximum. The maximum of the probability is obtained efficiently by iteration of the Newton-Raphson method:

$$(\ln Z_a)_{n+1} = (\ln Z_a)_n - \gamma \sum_b (H_{ab}^{-1})_n (F_b)_n, \qquad (9)$$

with $(\ln Z_a)_n$ the values of the partition functions at iteration n, and γ a scaling factor that limits the maximum step size and increases the radius of convergence.

The limit of the sequence of Eq. (9) gives the maximum likelihood estimate of the logarithms of the partition functions of all the systems and thus the maximum likelihood estimate of their free energies. The maximum likelihood estimate is asymptotically unbiased with the constraint that the free energies of the system cannot become infinite [12] (which is always the case in practice). Furthermore, the current estimator asymptotically has the minimum variance since the third order derivatives of the log likelihood are finite for any values of the work or the parameters [12]. Thus, for the limit of a large data set the present analysis

gives the optimal asymptotically unbiased estimate of the free energies.

For $N = 2$ systems the equation $F_a = 0$ formally equals that of the Bennett acceptance ratio method, and thus, as Bennett showed [4], it converges asymptotically to the free energy perturbation method [10] (or the Jarzynski equality [5]) in the limit of equilibrium sampling from only one system.

Example.—We illustrate the method by applying it to the analysis of three parallel tempering simulations [13,14] corresponding to alchemical mutations in vacuum between pairs of the capped amino acids [18] valine (V), threonine (T), and asparagine (N) as shown in Fig. 1; capped amino acids are widely used models in biophysical studies. Here, we determine the room temperature free energy differences $\Delta F_{V \to T}$ and $\Delta F_{V \to N}$, and their variances as a function of increasing the number of pathways to demonstrate the power of the current approach.

We sample the canonical ensembles of each amino acid using a molecular-dynamics–based parallel tempering algorithm [13,14] that readily equilibrates each state despite the high rotational barriers of the χ_1 angles [19,20]. Parallel tempering is performed with 12 heat baths $T_0 = 255$ K, $T_1 = 300$ K, $T_{n+1} = 300 + (n \times 50)$ K, $n \in$
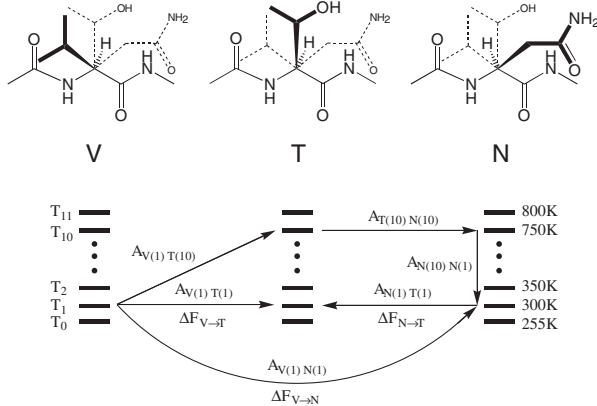

FIG. 1. Sampled thermodynamic states for capped amino acids valine (V), threonine (T), and asparagine (N). The amino acids are modeled with the CHARMM energy function [18,23] using a parametric potential: $V(\lambda_V, \lambda_T, \lambda_N) = V^b + V_{EE}^{nb} + \sum_X^{V,T,N} V_{XX}^{vdw} + \sum_X^{V,T,N} \lambda_X (V_{XX}^{elec} + V_{XE}^{nb})$, with λ_V, λ_T, and λ_N the parameters that scale the nonbonded and electrostatic energies of the amino acid of interest, V^b the bonded energy terms, V_{EE}^{nb} the nonbonded energy terms of the environment E (backbone), V_{XX}^{vdw} and V_{XX}^{elec} the van der Waals and electrostatic energy terms within the side chain of amino acid X (from the set $\{V, T, N\}$), and V_{XE}^{nb} the nonbonded energy terms of side chain X with the environment. We implement the potential with the BLOCK facility [24] of the CHARMM program [18]. The bold side chains have $\lambda_X = 1$, the dashed ones have $\lambda_X = 0$; for example, the state labeled T has $(\lambda_V, \lambda_T, \lambda_N) = (0, 1, 0)$. The lower part of the picture shows several thermodynamic states from the parallel tempering simulation. The arrows show a small subset of all the possible switching pathways that contribute to the evaluation of the 300 K free energy differences.

$\{1, 2, \ldots, 10\}$; the temperatures are controlled via Langevin dynamics with a friction coefficient of 40 ps^{-1}. The integration time step is 2 fs; SHAKE constraints [21] are applied for bonds involving hydrogen atoms. Every 10 ps we attempt to swap either all the baths (T_{2i}, T_{2i+1}) or all the baths (T_{2i-1}, T_{2i}), with i integer. The length of each simulation is 360 ns (36 000 swap attempts). We save snapshots every 2 ps and obtain equilibrated ensembles of 160 000 snapshots for each heat bath taken from the last 320 ns of the simulation.

Figure 2 shows several scatterplots of the simultaneous estimates of the 300 K free energy differences from equilibrium samples at N_T temperatures, using N_p data points for each transition. The centroid of the $N_T = 12$, $N_p = 5000$ set is close to the centroid of each scatterplot, which shows that the estimate is already unbiased for $N_p = 500$. The estimates of the free energy at 300 K from the coordinates of this centroid are $\Delta F_{V \to T} = -16.33 \pm 0.01$ kcal/mol and $\Delta F_{V \to N} = -73.78 \pm 0.01$ kcal/mol; the values of the free energy are lower for the larger groups (see Fig. 1), essentially due to their more negative energies.

A summary of the scatterplots of Fig. 2 and of other types of analysis of the sampled data is presented in Table I. The free energy perturbation method [10] is systematically biased for the small samples [22] and gives incorrect results even when the complete data set is used. The average of the forward and backward data [22] is
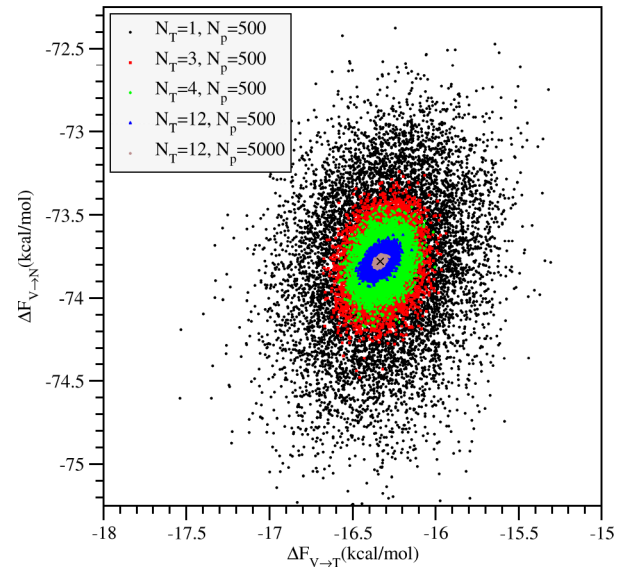

FIG. 2 (color online). Scatterplots of the 300 K free energy differences for the V $\to$ T and V $\to$ N transitions (see Fig. 1). Each scatterplot is build from 10 000 estimates of the free energy differences using random subsets of the work data. These subsets have length N_p for all possible pairs of the $3 \times N_T$ states of the three dipeptides and N_T heat baths. The $N_T = 1$ set contains only the 300 K ensembles of V, T, and N; the sets with more than one temperature contain the ensembles of the N_T lowest temperature heat baths. The symbol (X) marks the centroid of the set with the least scatter.

TABLE I. Table of the 300 K free energy estimates in kcal/mol from equilibrium samples of N_s peptides at N_T temperatures, using N_p data points for each transition. The standard deviations σ were calculated from 10 000 random samples; those in parentheses are analytical estimates from one sample [4].

N_s	N_T	N_p	$\langle \Delta F_{V \to T} \rangle$	$\sigma_{V \to T}$	$\langle \Delta F_{V \to N} \rangle$	$\sigma_{V \to N}$
1[a]	1	500	−14.22	0.85	−70.68	0.86
1[a]	1	1.6×10^5	−16.86		−71.72	
1[b]	1	500	−18.25	0.85	−76.59	1.23
1[b]	1	1.6×10^5	−17.426		−74.90	
2[c]	1	500	−16.23	0.60	−73.63	0.75
2[c]	1	1.6×10^5	−17.14		−73.31	
2[d]	1	500	−16.31	0.31	−73.81	0.51
2[d]	1	5000	−16.30	0.10	−73.78	0.16
2[d]	1	1.6×10^5	−16.30	(0.02)	−73.78	(0.04)
3[e]	1	500	−16.31	0.30	−73.80	0.44
3[e]	3	500	−16.32	0.11	−73.79	0.16
3[e]	4	500	−16.32	0.08	−73.79	0.10
3[e]	8	500	−16.33	0.04	−73.78	0.05
3[e]	12	500	−16.33	0.04	−73.78	0.04
3[e]	1	5000	−16.30	0.09	−73.78	0.13
3[e]	3	5000	−16.32	0.03	−73.79	0.05
3[e]	4	5000	−16.32	0.02	−73.79	0.03
3[e]	8,12	5000	−16.33	0.01	−73.78	0.01

[a]Free energy perturbation, initial.
[b]Free energy perturbation, final.
[c]Free energy perturbation, average.
[d]Bennett's acceptance ratio.
[e]Multistate acceptance ratio.

somewhat better but still considerably less accurate than the Bennett acceptance ratio analysis of the same data set. Inclusion of more pathways in the multistate acceptance ratio analysis improves the statistics and keeps the estimate consistent. The striking feature of this table is the scaling of the multistate acceptance ratio standard deviation σ; it is approximately proportional to $N_T^{-1} N_p^{-1/2}$ for a wide range of temperatures.

To estimate the efficiency of this method, we used random subsets of the microstates instead of the random subsets of the work used for Fig. 2 and Table I. We picked 500 random structures from each thermodynamic state to create the arrays of the work for all pathways; these work data are correlated. The estimate of $\Delta F_{V \to N}$ from 10 000 repetitions of this procedure is $\Delta F_{V \to N} = -73.78 \pm 0.10$ kcal/mol. This corresponds to a reduction of the standard deviation by a factor of 5.1 compared to that of the Bennett acceptance ratio between V and N at 300 K shown in Table I (entry 2[d], 1, 500); thus, the analysis of 1 ns of total simulation obtains the same accuracy as 26 ns analyzed with the Bennett method between the two end states.

Conclusions.—We have presented the asymptotically optimal way to obtain the free energy differences from samples of the work data between multiple states. The resulting multistate acceptance ratio method is numerically efficient and stable. We have demonstrated the applicability of this approach to the analysis of parallel tempering simulations and have shown that it provides estimates of the resulting free energy differences that are more precise and accurate than those from the Bennett method between two states. We are applying the approach to a range of problems (e.g., relative solvation free energies of amino acids). The method can be used also to enhance multi-ligand binding simulations or the analysis of experimental work data [2,3], as an extension of the approach of Hummer and Szabo [8].

[1] T. Simonson, G. Archontis, and M. Karplus, Acc. Chem. Res. **35**, 430 (2002).
[2] D. Collin *et al.*, Nature (London) **437**, 231 (2005).
[3] C. Cecconi *et al.*, Science **309**, 2057 (2005).
[4] C. H. Bennett, J. Comput. Phys. **22**, 245 (1976).
[5] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[6] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[7] G. E. Crooks, Phys. Rev. E **61**, 2361 (2000).
[8] G. Hummer and A. Szabo, Proc. Natl. Acad. Sci. U.S.A. **98**, 3658 (2001).
[9] M. R. Shirts *et al.*, Phys. Rev. Lett. **91**, 140601 (2003).
[10] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954).
[11] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).
[12] E. Lehmann and C. G. Casella, *Theory of Point Estimation* (Springer, New York, 1998), 2nd ed.
[13] C. J. Geyer, in *Proceedings of the 23rd Symposium on the Interface*, edited by E. M. Keramidas, Computing Science and Statistics (Interface Foundation, Fairfax Station, VA, 1991), pp. 156–163.
[14] U. H. E. Hansmann, Chem. Phys. Lett. **281**, 140 (1997).
[15] D. J. Evans and D. J. Searles, Adv. Phys. **51**, 1529 (2002).
[16] J. A. Anderson, Biometrika **59**, 19 (1972).
[17] T. Bayes, Phil. Trans. R. Soc. London **53**, 370 (1763).
[18] B. R. Brooks *et al.*, J. Comput. Chem. **4**, 187 (1983).
[19] T. Straatsma and J. McCammon, J. Chem. Phys. **95**, 1175 (1991).
[20] C. Bartels and M. Karplus, J. Comput. Chem. **18**, 1450 (1997).
[21] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, J. Comput. Phys. **23**, 327 (1977).
[22] R. H. Wood *et al.*, J. Phys. Chem. **95**, 6670 (1991).
[23] A. D. MacKerell *et al.*, J. Phys. Chem. B **102**, 3586 (1998).
[24] B. Tidor, Ph.D. thesis, Harvard University, 1990.