Enhanced Sampling
&
Free Energy Calculations

Methods Based on Equilibrium and Non-equilibrium Simulations

Ye Mei [1]
Pengfei Li [2]

December 12, 2019

[1]samuel.y.mei@gmail.com
[2]lipengfei_mail@126.com

Dedicated to
Dr. Bernard Brooks and Dr. Gerhard König.

Contents

List of Figures

List of Tables

Preface

Should we type some words here? Maybe not, 'coz we are not talkative persons.

About the companion website

The website[1] for this file contains:

- A link to (freely downloadable) latest version of this document.
- Link to some implementations of WHAM.
- Other stuff might appear in the near future (HOPEFULLY!).

Acknowledgements

- YM wants to express his special thanks to Dr. Bernard Brooks[2] and Dr. Gerhard König for helping him toddle in this field.
- We'll also like to thank Dr. Xiangyu Jia[3], Dr. Meiting Wang, Dr. Wei Liu and Dr. Fengjiao Liu for many helpful discussions.

Ye Mei
State Key Laboratory of Precision Spectroscopy
East China Normal University
Shanghai 200062 China
`https://qclassic.wordpress.com/`

[1] `https://github.com/samuelymei/`
[2] `https://www.lobos.nih.gov/cbs/`
[3] `https://research.shanghai.nyu.edu/centers-and-institutes/chemistry/people/xiangyu-jia`

1

Introduction

"Everything should be made as simple as possible but not simpler."
– Albert Einstein,

Computer simulations of biological systems have made much progress in the past decades. A battery of methods at different levels of sophistication and complexity have been proposed.

However, we are still facing grand difficulties from three aspects, i.e. accuracy of Hamiltonians, efficiency of sampling and reliability of postprocessing methods.[1]

In this booklet, we will not cover the whole spectrum of methods for enhanced samplings and free energy calculations, but only summarize some basic ideas. More complicated implementations of these methods, for instance 2-dimensional replica exchange molecular dynamics simulations, will not be discussed.

Recently, there is one special issue focusing on the methodologies of free energy calculations on Journal of Chemical Theory and Computation (Free Energy Calculations: Three Decades of Adventure in Chemistry and Biophysics, Journal of Chemical Theory and Computation, Volume 10, Issue 7, 2014, `https://pubs.acs.org/toc/jctcce/10/7`). There is also a special issue on the recent development in enhanced sampling methods for molecular systems (Special Topic on Enhanced Sampling for Molecular Systems, Journal of Chemical Physics, Volume 149, Issue 7, 2018, `https://aip.scitation.org/toc/jcp/149/7`).

There are also some good papers for reference

- Andrew Pohorille, Christopher Jarzynski and Christophe Chipot, Good Practices in Free-Energy Calculations, Journal of Physical Chemistry B, 2010, 114 (32), 10235–10253

- Daniel M. Zuckerman, Equilibrium Sampling in Biomolecular Simulations, Annual Review of Biophysics, 2011, 40:41–62

There are also two books on this topic you might be interested in:

- Free Energy Calculations: Theory and Applications in Chemistry and Biology, Editors: Christophe Chipot, Andrew Pohorille, ISBN 978-3-540-38448-9, Springer-Verlag Berlin Heidelberg, 2007
- Free Energy Computations: A Mathematical Perspective, Author: Tony Lelievre, Gabriel Stoltz, Mathias Rousset, ISBN-13: 978-1848162471, Imperial College Press, 2010

Before we move into the major content of this booklet, we would like to review some fundamentals that underlie the methods introduced in the following chapters. The first one is the canonical partition function Q for Hamiltonian $H(\mathbf{x}, \mathbf{p}_x)$, which is defined as

$$\begin{aligned} Q(N,V,T) =& \frac{1}{h^{3N}N!}\iint \exp[-\beta H(\mathbf{x},\mathbf{p}_x)]\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_\mathbf{x} \\ =& \frac{1}{\Lambda^{3N}N!}Z(N,V,T), \end{aligned} \tag{1.0.0.1}$$

where $\mathbf{x}$ and $\mathbf{p}_x$ are the coordinates and the conjugate momenta, respectively,

$$Z(N,V,T) = \int \exp\left(-\beta U(\mathbf{x})\right)\mathrm{d}\mathbf{x} \tag{1.0.0.2}$$

is the configurational integral, Λ is the temperature-dependent de Broglie wavelength, and $U(\mathbf{x})$ is the potential energy.

The partition function Q can also be defined in energy space as

$$Q(N,V,T) = \int \exp\left(-\beta E\right)\Omega_{tot}(N,V,E)\,\mathrm{d}E, \tag{1.0.0.3}$$

where

$$\Omega_{tot}(N,V,E) = \frac{1}{h^{3N}N!}\iint_{V^N}\delta(H(\mathbf{x},\mathbf{p}_x)-E)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_\mathbf{x} \tag{1.0.0.4}$$

is the complete density of states. Correspondingly, we can also define the configurational density of states as

$$\Omega_{con}(E) \propto \frac{1}{N!}\int_{V^N}\delta(U(\mathbf{x})-E)\,\mathrm{d}\mathbf{x}. \tag{1.0.0.5}$$

The Helmholtz free energy is defined in terms of the canonical partition function as

$$A = -\beta^{-1}\ln Q(N,V,T), \tag{1.0.0.6}$$

which connects thermodynamics and statistical mechanics. If we can estimate the value of Q, we can calculate A. However, evaluating Q is very difficult in most cases. Fortunately, we are only interested in the free energy

differences, ΔA, between two systems or two states of a system denoted by 0 and 1, respectively

$$\Delta A = -\beta^{-1} \ln Q_1/Q_0. \tag{1.0.0.7}$$

For most cases we are dealing with, the masses of particles in systems 0 and 1 are the same, Eq. 1.0.0.7 can be rewritten in terms of the configurational integrals Z_0 and Z_1

$$\Delta A = -\beta^{-1} \ln Z_1/Z_0. \tag{1.0.0.8}$$

In the following chapters, the systems 0 and 1 may differ in several ways. They may have different Hamiltonians, H_0 and H_1.

$$Q_0 = \frac{1}{N!h^{3N}} \int \exp\left[-\beta H_0(\mathbf{x}, \mathbf{p}_x)\right] \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x \tag{1.0.0.9}$$

$$Q_1 = \frac{1}{N!h^{3N}} \int \exp\left[-\beta H_1(\mathbf{x}, \mathbf{p}_x)\right] \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x \tag{1.0.0.10}$$

Or they may be characterized by different values of a macroscopic parameter, such as temperature.

$$Q_0 = \frac{1}{N!h^{3N}} \int \exp\left[-\beta_0 H(\mathbf{x}, \mathbf{p}_x)\right] \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x \tag{1.0.0.11}$$

$$Q_1 = \frac{1}{N!h^{3N}} \int \exp\left[-\beta_1 H(\mathbf{x}, \mathbf{p}_x)\right] \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x \tag{1.0.0.12}$$

Finally, they may correspond to different regions in the phase space accessible to the system

$$Q_0 = \frac{1}{N!h^{3N}} \int_{\Gamma_0} \exp\left[-\beta H(\mathbf{x}, \mathbf{p}_x)\right] \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x \tag{1.0.0.13}$$

$$Q_1 = \frac{1}{N!h^{3N}} \int_{\Gamma_1} \exp\left[-\beta H(\mathbf{x}, \mathbf{p}_x)\right] \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x \tag{1.0.0.14}$$

where Γ_0 and Γ_1 may refer to different conformations of a flexible molecules, or the bound and unbound structures of a protein-ligand complex, etc.

In canonical ensemble (with NVT fixed), the probability of a microstate is

$$\rho(\mathbf{x}) = \frac{1}{Z} \exp\left(-\beta E(\mathbf{x})\right), \tag{1.0.0.15}$$

where $E(\mathbf{x})$ is the potential energy of this microstate. With this probability we can calculate the expectation of any operator $\hat{O}$ on configurations via

$$\langle O \rangle = \frac{\int \hat{O}(\mathbf{x}) \exp\left(-\beta E(\mathbf{x})\right) \mathrm{d}\mathbf{x}}{Z}. \tag{1.0.0.16}$$

Besides state free energy, we may also be interested in free energy profiles

along one or several degrees of freedom

$$\begin{aligned} A(\xi) =& -\beta^{-1}\ln Z(\xi) \\ =& -\beta^{-1}\ln\int \exp\left(-\beta U\right)\delta(\xi-\xi(\mathbf{x}))\,\mathrm{d}\mathbf{x} \\ =& -\beta^{-1}\ln\int \exp\left(-\beta U(\xi(\mathbf{x})=\xi)\right)|\mathbf{J}|\,\mathrm{d}q_1\cdots\mathrm{d}q_{N-1}, \end{aligned} \tag{1.0.0.17}$$

where $\mathbf{J}$ is the Jacobian matrix upon changing from Cartesian to generalized coordinates with its element defined as $[\mathbf{J}(\mathbf{q})]_{ij} = \partial x_i/\partial q_j$ with $q_N = \xi$. $|\mathbf{J}|$ is its determinant. Its gradient over ξ is

$$\begin{aligned} \frac{\partial A(\xi)}{\partial\xi} =& -\beta^{-1}\frac{\int\frac{\partial}{\partial\xi}\left(e^{-\beta U}|\mathbf{J}|\right)\mathrm{d}q_1\cdots\mathrm{d}q_{N-1}}{\int e^{-\beta U}\delta(\xi-\xi(\mathbf{x}))\,\mathrm{d}\mathbf{x}} \\ =&\frac{\int e^{-\beta U}\left[\frac{\partial U}{\partial\xi}-\beta^{-1}\frac{1}{|\mathbf{J}|}\frac{\partial|\mathbf{J}|}{\partial\xi}\right]|\mathbf{J}|\,\mathrm{d}q_1\cdots\mathrm{d}q_{N-1}}{\int e^{-\beta U}\delta(\xi-\xi(\mathbf{x}))\,\mathrm{d}\mathbf{x}} \\ =&\frac{\int e^{-\beta U}\left[\frac{\partial U}{\partial\xi}-\beta^{-1}\frac{\partial\ln|\mathbf{J}|}{\partial\xi}\right]|\mathbf{J}|\,\mathrm{d}q_1\cdots\mathrm{d}q_{N-1}}{\int e^{-\beta U}\delta(\xi-\xi(\mathbf{x}))\,\mathrm{d}\mathbf{x}} \\ =&\frac{\int e^{-\beta U}\left[\frac{\partial U}{\partial\xi}-\beta^{-1}\frac{\partial\ln|\mathbf{J}|}{\partial\xi}\right]\delta(\xi-\xi(\mathbf{x}))\,\mathrm{d}\mathbf{x}}{\int e^{-\beta U}\delta(\xi-\xi(\mathbf{x}))\,\mathrm{d}\mathbf{x}} \\ =&\left\langle\frac{\partial U}{\partial\xi}-\beta^{-1}\frac{\partial\ln|\mathbf{J}|}{\partial\xi}\right\rangle_{\xi}. \end{aligned} \tag{1.0.0.18}$$

Here, $-\frac{\partial U}{\partial\xi}+\beta^{-1}\frac{\partial\ln|\mathbf{J}|}{\partial\xi}$ is the generalized force on ξ to be averaged over the degrees of freedom other than ξ itself. Therefore, $A(\xi)$ is called the potential of mean force. *Note: Some define the potential of mean force as* $\left\langle\frac{\partial U}{\partial\xi}\right\rangle_{\xi}$ *only. But we do not strictly differentiate potential of mean force and free energy profile here.*

Let us take protein-ligand binding

$$\mathrm{P} + \mathrm{L} \rightleftharpoons \mathrm{P\text{-}L}$$

as an example to illustrate how the simulations and free energy methods are used in real problems. The equilibrium constant, K_b, is defined as

$$K_b = \frac{[\mathrm{P\text{-}L}]}{[\mathrm{P}][\mathrm{L}]}, \tag{1.0.0.19}$$

where [P-L], [P] and [L] are the equilibrium concentrations of the complex, protein and ligand, respectively. A standard binding free energy can be calculated via $\Delta G_{bind} \equiv -\beta^{-1}\ln\left[K_b C^0\right]$, where C^0 is the standard state concentration of 1 mol/liter ($\equiv 1/1661\text{Å}^3$). K_b can be expressed in terms of a ratio of configurational integrals as

$$K_b = \frac{1}{[L]}\frac{N\int_{site}\mathrm{d}(\mathbf{1})\int_{bulk}\mathrm{d}(\mathbf{2})\cdots\int_{bulk}\mathrm{d}(\mathbf{N})\int\mathrm{d}\mathbf{X}e^{-\beta U}}{\int_{bulk}\mathrm{d}(\mathbf{1})\int_{bulk}\mathrm{d}(\mathbf{2})\cdots\int_{bulk}\mathrm{d}(\mathbf{N})\int\mathrm{d}\mathbf{X}e^{-\beta U}}, \tag{1.0.0.20}$$

where U is the total potential energy of the system, $(\mathbf{1})$, $(\mathbf{2})$, $\cdots$, $(\mathbf{N})$ and $\mathbf{X}$ are the coordinates of the N ligand molecules and the remaining atoms, respectively. For simplicity, we omit the integrals over the $(N-1)$ ligands in bulk, and we notice that $\int_{bulk} \mathrm{d}(\mathbf{1}) = V_{bulk} \int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r_1} - \mathbf{r}^*)$. Then, we have

$$\begin{aligned} K_b =& \frac{1}{[L]} \frac{N \int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta U}}{V_{bulk} \int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r_1} - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta U}} \\ =& \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta U}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r_1} - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta U}}. \end{aligned} \tag{1.0.0.21}$$

A direct calculation of this ratio is not easy. Practically, we can define a series of intermediate states. Thereupon, the calculation of this ratio can be facilitated by

$$\begin{aligned} K_b = \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta U}}{Z_1} \times \frac{Z_1}{Z_2} \times \cdots \times \frac{Z_{n-1}}{Z_n} \times \\ \frac{Z_n}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r_1} - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta U}} \end{aligned} \tag{1.0.0.22}$$

There are two categories of methods for computing K_b, i.e. the alchemical strategy[2] and the PMF-based strategy[3]. A comparison of these two strategies can be found in Ref. [4] and [5]. For each ratio in Eq. 1.0.0.22, we shall design a proper simulation and employ a suitable free energy method to calculate the free energy associated with it. Enhanced sampling might be necessary for convergence.

In alchemical strategy, the ligand is decoupled from its environment in the binding pocket and then appears in water. However, a series of steps with restraints on the conformation, translation and rotation are introduced.

Equation 1.0.0.22 is now realized as

$$
\begin{aligned}
K_b = & \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta U_1}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c]}} \times \\
& \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c]}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c+U_t]}} \times \\
& \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c+U_t]}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c+U_t+U_r]}} \times \\
& \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c+U_t+U_r]}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_0+U_c+U_t+U_r]}} \times \\
& \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_0+U_c+U_t+U_r]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_0+U_c+U_t]}} \times \\
& \frac{\int_{bulk} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_0+U_c+U_t]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U_0+U_c]}} \times \\
& \frac{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U_0+U_c]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c]}} \times \\
& \frac{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta U_1}}.
\end{aligned}
\tag{1.0.0.23}
$$

In PMF-based strategy, the ligand is gradually pulled out of the binding pocket into water. Similarly, restraints on the conformation, translation and rotation should also be applied when pulling the ligand molecule. Equation 1.0.0.22 is now realized as

$$
\begin{aligned}
K_b = & \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta U}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c]}} \times \\
& \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c]}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c+U_o]}} \times \\
& \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c+U_o]}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c+U_o+U_a]}} \times \\
& \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c+U_o+U_a]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathrm{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c+U_o]}} \times \\
& \frac{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathrm{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c+U_o]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathrm{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c]}} \times \\
& \frac{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathrm{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathrm{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta U}}.
\end{aligned}
\tag{1.0.0.24}
$$

It is worth emphasizing that this PMF-based strategy does not yield the real entry/escaping pathways nor the free energy barriers.

2

Enhanced Sampling

"Keep the smart guys around you."

– Bernard R. Brooks

From the definition, the free energy of a specific system is dominated by phase space regions with a low potential energy. However, these regions might be separated by high energy barriers. Transitions among these potential energy wells are often hindered by these barriers. According to the Boltzmann's Law, the probability of a sample $\mathbf{R}$ being visited is proportional to the Boltzmann's factor $\exp\left[-\beta E(\mathbf{R})\right]$, where $\beta = 1/k_B T$ is called the inverse temperature. k_B is the Boltzmann constant and T is the temperature. According to some experience, in a $100\,ns$ simulation, the system can overcome a barrier of $10\,k_B T$, which is $6\,kcal/mol$ at room temperature ($300\,K$). If the barrier is $1.5\,kcal/mol$ higher, it takes about $1\,\mu s$ (10 times longer) in average for the system to go over the barrier. If the barrier height reaches $9\,kcal/mol$, it takes $10\,\mu s$. And so on. With modern computers, the longest all-atom molecular dynamics simulation for biological systems is probably the one done by D.E. Shaw, which was on a time scale of 1 ms on a special-purpose computer "Anton". For most classical molecular dynamics simulations, the time scales are normally several μs to tens of μs. For simulations using expensive Hamiltonians, such as in QM/MM simulations, the time scales that can be reached are usually three orders shorter. Clearly, molecular dynamics simulations are plagued by a timescale problem. In order to observe abundant transitions among these energy minima, which is required by free energy calculations, enhanced samplings are often indispensable. As shown in the Boltzmann's factor, the essential quantity that determines the rate of transitions is βE. In order to accelerate the phase space sampling, we can either increase the temperature or decrease the energy barrier. All the methods shown below can be classified into these two categories.

2.1 Replica Exchange Molecular Dynamics

2.1.1 Temperature-Replica Exchange Molecular Dynamics

Temperature replica exchange molecular dynamics (T-REMD) is one class of parallel tempering methods developed by Hansmann, Okamoto and Sugita[6, 7, 8] based on many ideas in a category of methods called *generalized-ensemble algorithm*. It is an extension of the well-known simulated annealing method. The basic idea of REMD is schematically summarized in Fig. 2.1. In REMD, the system is replicated into **M** *non-interacting* copies (replicas). Each replica is coupled to a bath at temperature T_m, $(m = 1, \ldots, M)$. At a certain time, the system is at state X, which can be denoted as $X = \left(x_1^{[i(1)]}, \ldots, x_M^{[i(M)]}\right) = \left(x_{m(1)}^{[1]}, \ldots, x_{m(M)}^{[M]}\right)$. Here, we used i and m to label the replica and the temperature respectively. Because the replicas are non-interacting, the weight-factor for a state X in this generalized ensemble is a direct product of the Boltzmann factors for each replica, i.e.

$$W_{REM}(X) = \prod_{m=1}^{M} \exp\left(-\beta_m H\left(q^{[i(m)]}, p^{[i(m)]}\right)\right) = \prod_{i=1}^{M} \exp\left(-\beta_{m(i)} H\left(q^{[i]}, p^{[i]}\right)\right) \tag{2.1.1.1}$$

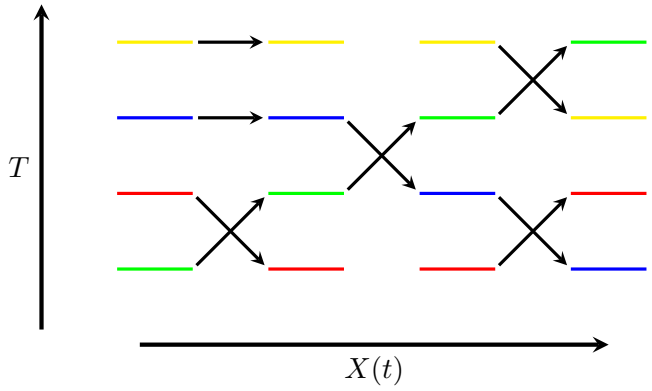

Figure 2.1: A schematic representation of replica exchange molecular dynamics.

Now, we exchange the temperatures of a pair of replicas

$$\begin{cases} x_m^{[i]} \equiv \left(q^{[i]}, p^{[i]}\right)_m \Rightarrow x_n^{[i]'} \equiv \left(q^{[i]}, p^{[i]'}\right)_n \\ x_n^{[j]} \equiv \left(q^{[j]}, p^{[j]}\right)_n \Rightarrow x_m^{[j]'} \equiv \left(q^{[j]}, p^{[j]'}\right)_m \end{cases}, \tag{2.1.1.2}$$

where

$$\begin{cases} p^{[i]'} \equiv \sqrt{\frac{T_n}{T_m}} p^{[i]} \\ p^{[j]'} \equiv \sqrt{\frac{T_m}{T_n}} p^{[j]} \end{cases} . \tag{2.1.1.3}$$

The exchange rule is not trivial. In order for this exchange process to converge towards an equilibrium distribution, it is sufficient to impose the detailed balance condition on the transition probability $w(X \to X')$:

$$W_{REM}(X) w(X \to X') = W_{REM}(X') w(X' \to X). \tag{2.1.1.4}$$

Then we have

$$\begin{aligned} & \frac{w\left(X \to X'\right)}{w\left(X' \to X\right)} \\ = & \frac{W_{REM}(X')}{W_{REM}(X)} \\ = & \frac{\exp\left(-\beta_m H\left(q^{[j]}, p^{[j]'}\right)\right) \exp\left(-\beta_n H\left(q^{[i]}, p^{[i]'}\right)\right)}{\exp\left(-\beta_m H\left(q^{[i]}, p^{[i]}\right)\right) \exp\left(-\beta_n H\left(q^{[j]}, p^{[j]}\right)\right)} \\ = & \frac{\exp\left\{-\beta_m\left[K\left(p^{[j]'}\right) + U\left(q^{[j]}\right)\right] - \beta_n\left[K\left(p^{[i]'}\right) + U\left(q^{[i]}\right)\right]\right\}}{\exp\left\{-\beta_m\left[K\left(p^{[i]}\right) + U\left(q^{[i]}\right)\right] - \beta_n\left[K\left(p^{[j]}\right) + U\left(q^{[j]}\right)\right]\right\}} \\ = & \frac{\exp\left\{-\beta_m\left[\frac{T_m}{T_n} K\left(p^{[j]}\right) + U\left(q^{[j]}\right)\right] - \beta_n\left[\frac{T_n}{T_m} K\left(p^{[i]}\right) + U\left(q^{[i]}\right)\right]\right\}}{\exp\left\{-\beta_m\left[K\left(p^{[i]}\right) + U\left(q^{[i]}\right)\right] - \beta_n\left[K\left(p^{[j]}\right) + U\left(q^{[j]}\right)\right]\right\}} \\ = & \frac{\exp\left\{-\beta_n K\left(p^{[j]}\right) - \beta_m K\left(p^{[i]}\right)\right\}}{\exp\left\{-\beta_m K\left(p^{[i]}\right) - \beta_n K\left(p^{[j]}\right)\right\}} \frac{\exp\left\{-\beta_m U\left(q^{[j]}\right) - \beta_n U\left(q^{[i]}\right)\right\}}{\exp\left\{-\beta_m U\left(q^{[i]}\right) - \beta_n U\left(q^{[j]}\right)\right\}} \\ = & \exp\{-\Delta\}. \end{aligned} \tag{2.1.1.5}$$

where $\Delta = \left[\beta_n - \beta_m\right]\left[U\left(q^{[i]}\right) - U\left(q^{[j]}\right)\right]$. It can be seen that the kinetic energy terms are fully canceled out. This can be satisfied by the usual Metropolis criterion:

$$w\left(X \to X'\right) \equiv w\left(x_m^{[i]} \middle| x_n^{[j]}\right) = \begin{cases} 1, & \text{if } \Delta \leq 0 \\ \exp\left(-\Delta\right), & \text{if } \Delta > 0 \end{cases} \tag{2.1.1.6}$$

After long time simulations, all the replicas have arrived at a global equilibrium. In order to calculate the the free energy or the ensemble average of an operator $\hat{A}$ at T_m, we can extract all the snapshots that have a temperature T_m from M trajectories, if this temperature was among the M chosen temperatures. However, the optimal way is to use Weighted Histogram Analysis Method in Section 3.1.4 or the Multistate Bennett Acceptance Ratio method in Section 3.1.5.

2.1.2 Hamiltonian-Replica Exchange Molecular Dynamics

Another type of REMD simulation is called Hamiltonian replica exchange molecular dynamics (H-REMD), in which each replicas has its own Hamiltonian, but is coupled to the same temperature. One example is the H-REMD simulation for a torsional angle. The mth replica has a torsional energy term of

$$H_m(\phi) = \lambda(m) \sum_n (V_n/2) \left(1 + \cos\left[n\phi - \delta\right]\right), \tag{2.1.2.1}$$

where λ is a control parameter. $\lambda(0) = 1$ corresponds to the unbiased state and at $\lambda(M)$ (usually $\lambda(M) = 0$) the torsional motion of this dihedral angle has a smaller barrier.

Another example of HREMD is pH-REMD, in which each replica is coupled with different pH of the solution. In other words, the chemical potential of hydronium in each replica is different . Therefore, the protonation states (or probability of being protonated or deprotonated) of titratable residues in each replica may differ from those in other replicas. In the simulations, the protonation states of titratable residues have their protonation states alternated according to the Metropolis criterion

$$P = \begin{cases} 1, & \text{if } \Delta G_{\mathrm{P_A} \to \mathrm{P_A H^+}} \leq 0 \\ \exp\left(-\beta \Delta G_{\mathrm{P_A} \to \mathrm{P_A H^+}}\right), & \text{if } \Delta G_{\mathrm{P_A} \to \mathrm{P_A H^+}} > 0 \end{cases} \tag{2.1.2.2}$$

using Monte Carlo. The derivation of $\Delta G_{\mathrm{P_A} \to \mathrm{P_A H^+}}$ is shown below.

Free energy of molecule A in solution with a concentration [A] can be written as

$$\Delta G_{\mathrm{A}} = \Delta G^0_{\mathrm{A}} + \beta^{-1} \ln \frac{[\mathrm{A}]}{C_0},$$

in which ΔG^0_{A} is the free energy of molecule A at the standard state C_0, i.e. 1 mol/L. The free energy change for a reaction

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{C}$$

can be written as

$$\Delta G = \Delta G_{\mathrm{C}} - \Delta G_{\mathrm{A}} - \Delta G_{\mathrm{B}} = \Delta G_0 + \beta^{-1} \ln \frac{[\mathrm{C}]\, C_0}{[\mathrm{A}]\, [\mathrm{B}]}.$$

At equilibrium, the free energy change is zero, we have

$$\Delta G_0 = -\beta^{-1} \ln \frac{[\mathrm{C}]\, C_0}{[\mathrm{A}]\, [\mathrm{B}]}, \tag{2.1.2.3}$$

in which $[\mathrm{A}]\,[\mathrm{B}] \,/\, [\mathrm{C}]\, C_0$ is called the dissociation constant K_a. So,

$$\Delta G_0 = \beta^{-1} \ln K_a. \tag{2.1.2.4}$$

Titration of a residue in a real protein can be written as

$$\mathrm{P_A} + \mathrm{H^+} \rightleftharpoons \mathrm{P_AH^+}$$

with

$$K_a = \frac{[\mathrm{P_A}]\,[\mathrm{H^+}]}{[\mathrm{P_AH^+}]\,C_0}$$

The fraction of the deprotonated species is calculated as

$$\begin{aligned} f_{[\mathrm{P_A}]} &= \frac{[\mathrm{P_A}]}{[\mathrm{P_A}] + [\mathrm{P_AH^+}]} \\ &= \frac{1}{1 + \frac{[\mathrm{P_AH^+}]}{[\mathrm{P_A}]}} \\ &= \frac{1}{1 + \frac{[\mathrm{P_A}][\mathrm{H^+}]}{C_0 K_a [\mathrm{P_A}]}} \\ &= \frac{1}{1 + \frac{1}{C_0 K_a}[\mathrm{H^+}]} \\ &= \frac{1}{1 + \frac{1}{K_a}10^{-\mathrm{pH}}} \end{aligned} \tag{2.1.2.5}$$

We can check the asymptotic behavior of this equation. At strong acidic condition ($\mathrm{pH} = -\infty$), $f_{[\mathrm{P_A}]} = 0$, indicating that the residue is 100 percent protonated. While at an extremely basic condition ($\mathrm{pH} = \infty$), $f_{[\mathrm{P_A}]} = 1$. This residue is 100 percent deprotonated. From the Henderson–Hasselbalch (HH) equation, the pK_a can be determined by the pH of the state when $[\mathrm{P_A}] \,/\, [\mathrm{P_AH^+}] = 1$ (the isoelectric point)

$$\begin{aligned} \mathrm{p}K_a &= -\log K_a \\ &= -\log \frac{[\mathrm{P_A}]}{[\mathrm{P_AH^+}]} - \log \frac{[\mathrm{H^+}]}{C_0} \\ &= -\log \frac{[\mathrm{P_A}]}{[\mathrm{P_AH^+}]} + \mathrm{pH}. \end{aligned} \tag{2.1.2.6}$$

The pK_a of each residue in a dipeptide has been determined by experiment. However, when this residue is located in a certain protein, its pK_a is different from that in the dipeptide. The difference is called the pK_a shift. Instead of measuring the pK_a for a residue in a protein, we are more interested in calculating/measuring the titration curve, which is the fraction of the deprotonated state as a function of pH. From Eq. 2.1.2.5, $f_{[P_A]}$ can be easily calculated if we know K_a or equivalently the standard free energy change of protonation in Eq. 2.1.2.4. The standard free energy can be calculated from the partition functions as

$$\begin{aligned}\Delta G_0 &= -\beta^{-1}\ln\frac{Q_{\mathrm{P_AH^+}}}{Q_{\mathrm{P_A}}Q_{\mathrm{H^+}}}\\ &= -\beta^{-1}\ln\frac{\iint\exp(-\beta E_{\mathrm{P_AH^+}})\,\mathrm{d}\mathbf{R}_H\,\mathrm{d}\mathbf{R}_o}{Q_{\mathrm{H^+}}\int\exp(-\beta E_{\mathrm{P_A}})\,\mathrm{d}\mathbf{R}_o},\end{aligned}$$

where $\mathbf{R}_H$ is the coordinates of the specific H atom and the other degrees-of-freedom (DoF) are denoted as $\mathbf{R}_o$. Generally, the absolute value of ΔG_0 is hardly computable. A relative protonation free energy $\Delta\Delta G$ is preferred and is more reliable. Theoretically, the reference state can be any state you like. But the protonation free energy of the dipeptide is often used. The reference protonation process can be written as

$$\mathrm{A} + \mathrm{H^+} \rightleftharpoons \mathrm{AH^+}$$

The free energy change from the reference state is

$$\begin{aligned}&\Delta\Delta G_0\\ =&\Delta G_0 - \Delta G_0^{ref}\\ =&-\beta^{-1}\ln\frac{\iint\exp(-\beta E_{\mathrm{P_AH^+}})\,\mathrm{d}\mathbf{R}_H\,\mathrm{d}\mathbf{R}_o}{Q_{\mathrm{H^+}}\int\exp(-\beta E_{\mathrm{P_A}})\,\mathrm{d}\mathbf{R}_o}\frac{Q_{\mathrm{H^+}}\int\exp(-\beta E_{\mathrm{A}})\,\mathrm{d}\mathbf{R}_o}{\iint\exp(-\beta E_{\mathrm{AH^+}})\,\mathrm{d}\mathbf{R}_H\,\mathrm{d}\mathbf{R}_o}\\ =&-\beta^{-1}\ln\frac{\iint\exp(-\beta E_{\mathrm{P_AH^+}})\,\mathrm{d}\mathbf{R}_H\,\mathrm{d}\mathbf{R}_o\int\exp(-\beta E_{\mathrm{A}})\,\mathrm{d}\mathbf{R}_o}{\int\exp(-\beta E_{\mathrm{P_A}})\,\mathrm{d}\mathbf{R}_o\iint\exp(-\beta E_{\mathrm{AH^+}})\,\mathrm{d}\mathbf{R}_H\,\mathrm{d}\mathbf{R}_o}\\ =&-\beta^{-1}\ln\frac{\iint\exp\left[-\beta\left(E^{bond}_{\mathrm{P_AH^+}}+E^{QM}_{\mathrm{P_AH^+}}+E^{ele}_{\mathrm{P_AH^+}}\right)\right]\mathrm{d}\mathbf{R}_H\exp\left(-\beta E^{other}_{\mathrm{P_AH^+}}\right)\mathrm{d}\mathbf{R}_o}{\iint\exp\left[-\beta\left(\mathrm{E}^{bond}_{\mathrm{AH^+}}+E^{QM}_{\mathrm{AH^+}}+E^{ele}_{\mathrm{AH^+}}\right)\right]\mathrm{d}\mathbf{R}_H\exp\left(-\beta E^{other}_{\mathrm{AH^+}}\right)\mathrm{d}\mathbf{R}_o}\\ &\cdot\frac{\int\exp\left(-\beta E_{\mathrm{A}}\right)\mathrm{d}\mathbf{R}_o}{\int\exp\left(-\beta E_{\mathrm{P_A}}\right)\mathrm{d}\mathbf{R}_o},\end{aligned}\tag{2.1.2.7}$$

where E^{bond} and E^{ele} are the bonded energy and electrostatic interaction energy related to this H atom, respectively. E^{QM} is the energy correction that *may* be required if the molecular mechanical Hamiltonian cannot well capture the energy of the system, such as the missing of charge transfer effect. The sum of the remaining energy term is denoted as E^{other}, which does not explicitly depend on the position of this specific H atom. Eq. 2.1.2.7 is not ready to be computed before some approximations are adopted.

First, we assume that the total energy can be well described by the MM Hamiltonians for both the state interested in and the reference state. Therefore,

$$E^{QM}_{\mathrm{P_AH^+}} = E^{QM}_{\mathrm{AH^+}} = Const,$$

and they can be removed from the integral.

Second, the bonded terms involving hydrogen atoms are usually constrained in the simulations. Therefore, the hydrogen atom in question has only one position and $E^{bond} = 0$. Now, the relative protonation free energy can be simplified as

$$\Delta\Delta G_0 = -\beta^{-1} \ln \frac{\int \exp\left(-\beta E^{ele}_{\mathrm{P_AH^+}}\right) \exp\left(-\beta E^{other}_{\mathrm{P_AH^+}}\right) \mathrm{d}\mathbf{R}_o}{\int \exp\left(-\beta E^{ele}_{\mathrm{AH^+}}\right) \exp\left(-\beta E^{other}_{\mathrm{AH^+}}\right) \mathrm{d}\mathbf{R}_o} \cdot \frac{\int \exp\left(-\beta E_{\mathrm{A}}\right) \mathrm{d}\mathbf{R}_o}{\int \exp\left(-\beta E_{P_{\mathrm{A}}}\right) \mathrm{d}\mathbf{R}_o}. \tag{2.1.2.8}$$

Note that $E_{\mathrm{A}} = E^{other}_{\mathrm{AH^+}}$ and $E_{\mathrm{P_A}} = E^{other}_{\mathrm{P_AH^+}}$, we have

$$\Delta\Delta G_0 = -\beta^{-1} \ln \frac{\int \exp\left(-\beta E^{ele}_{\mathrm{P_AH^+}}\right) \exp\left(-\beta E_{\mathrm{P_A}}\right) \mathrm{d}\mathbf{R}_o}{\int \exp\left(-\beta E_{\mathrm{P_A}}\right) \mathrm{d}\mathbf{R}_o} \tag{2.1.2.9}$$

$$\cdot \frac{\int \exp\left(-\beta E_{\mathrm{A}}\right) \mathrm{d}\mathbf{R}_o}{\int \exp\left(-\beta E^{ele}_{\mathrm{AH^+}}\right) \exp\left(-\beta E_{\mathrm{A}}\right) \mathrm{d}\mathbf{R}_o} \tag{2.1.2.10}$$

$$= -\beta^{-1} \ln \left\langle \exp\left(-\beta E^{ele}_{\mathrm{P_AH^+}}\right) \right\rangle_{\mathrm{P_A}} + \beta^{-1} \ln \left\langle \exp\left(-\beta E^{ele}_{\mathrm{AH^+}}\right) \right\rangle_{\mathrm{A}}$$

$$= \Delta G^{ele}_{\mathrm{P_AH^+}} - \Delta G^{ele}_{\mathrm{AH^+}} \tag{2.1.2.11}$$

Therefore,

$$-\beta^{-1} \ln 10 \cdot \mathrm{p}K_a = \Delta G^{ele}_{\mathrm{P_AH^+}} - \Delta G^{ele}_{\mathrm{AH^+}} - \beta^{-1} \ln 10 \cdot \mathrm{p}K^{ref}_a.$$

Using Eq. 2.1.2.6, at a certain pH the free energy difference between the deprotonated and the protonated state can be written as

$$\Delta G_{\mathrm{P_A} \to \mathrm{P_AH^+}} = \Delta G^{ele}_{\mathrm{P_AH^+}} + \beta^{-1}(\mathrm{pH} - \mathrm{p}K^{ref}_a) \ln 10 - \Delta G^{ele}_{\mathrm{AH^+}}.$$

In the above equation, $\Delta G^{ele}_{\mathrm{AH^+}}$ can be obtained from a free energy calculation of the model system by alchemically annihilation of the proton. However, $\Delta G^{ele}_{\mathrm{P_AH^+}}$ is unknown. Approximately, it can be replaced with $\Delta H^{ele}_{\mathrm{P_AH^+}}$ averaged over a few snapshots.[9] In order to accelerate the convergence, this pH-REMD is often coupled with other enhanced simulation methods, such as T-REMD[9] and EDS-REMD[10] (see section 2.10).

2.2 Simulated Tempering

Simulated tempering (ST) was proposed by Marinari and Parisi[11] and by Lyubartsev et al[12] in 1992, and by Geyer and Thompson[13] in 1995. In ST, there is only one trajectory with controlled jumps in temperature space, which is different from the implementation of T-REMD2.1.1, in which multiple trajectories are running in parallel. In ST, the simulation is carried out in an extended space defined by the configuration variables $\mathbf{X}$ and a new variable m. The latter can take M values ($m = 1, \ldots, M$). Corresponding to each m, there is an inverse temperature β_m. Let

$$\beta_0 > \beta_1 > \beta_2 > \cdots > \beta_m. \tag{2.2.0.1}$$

The probability distribution $\rho(\mathbf{X}, m)$ will be chosen to be

$$\rho(\mathbf{X}, m) \propto \exp\left[-H(\mathbf{X}, m)\right] \tag{2.2.0.2}$$

with

$$H(\mathbf{X}, m) \equiv \beta_m H(X) - g_m. \tag{2.2.0.3}$$

The total partition function for this extended ensemble is

$$\begin{aligned} Z &= \sum_{m=1}^{M} \int \mathrm{d}\mathbf{X} \exp\left[-H(\mathbf{X}, m)\right] \\ &= \sum_{m=1}^{M} \int \mathrm{d}\mathbf{X} \exp\left[-\beta_m H(X) + g_m\right] \\ &= \sum_{m=1}^{M} \exp\left(g_m\right) \int \mathrm{d}\mathbf{X} \exp\left[-\beta_m H(X)\right] \end{aligned} \tag{2.2.0.4}$$

For each β_m, there is a canonical ensemble with the probability for a configuration $\mathbf{X}$ follows the usual Boltzmann distribution, i.e.

$$\rho(X|m) \propto \exp\left(-\beta_m H(X)\right). \tag{2.2.0.5}$$

and the partition function (with $1/N!$ omitted)

$$Z_m = \int \mathrm{d}\mathbf{X} \exp\left[-\beta_m H(\mathbf{X})\right] = \exp\left(-\beta_m f_m\right). \tag{2.2.0.6}$$

where f_m is the associated free energy. With this definition of Z_m, the total partition function can be written as

$$Z = \sum_{m=1}^{M} \exp\left(g_m\right) Z_m, \tag{2.2.0.7}$$

where $\exp\left(g_m\right)$ can be thought as the weight for the mth canonical ensemble in this extended ensemble.

On the other hand, the probability of having a given value of m is simply given by

$$\rho_m = \frac{\int \mathrm{d}\mathbf{X} \exp\left[-H(\mathbf{X}, m)\right]}{Z} = \frac{Z_m \exp\left(g_m\right)}{Z} = \frac{1}{Z} \exp\left(-\beta_m f_m + g_m\right). \tag{2.2.0.8}$$

If we make the choice $g_m = \beta_m f_m$, then all the ρ_m become equal. However, f_m is usually unknown beforehand, or it may be never known.

The system may evolve in two types of steps: (1) usual displacements of particles at fixed temperature via molecule dynamics or Monte Carlo and (2) changes of reciprocal temperature with fixed positions of particles. In the first case, detailed balance can be easily maintained. In the second case, in order to maintain detailed balance

$$\rho(\mathbf{X}, m) P(\beta_m \to \beta_{m\pm1} | \mathbf{X}) = \rho(\mathbf{X}, m \pm 1) P(\beta_{m\pm1} \to \beta_m | \mathbf{X}), \tag{2.2.0.9}$$

transition takes place with the probability

$$P(\beta_m \to \beta_{m\pm1} | \mathbf{X}) = \min\left\{1, \exp\left[-(\beta_{m\pm1} - \beta_m) H(\mathbf{X}) + (g_{m\pm1} - g_m)\right]\right\} \tag{2.2.0.10}$$

following the Metropolis criteria.

2.3 Umbrella Sampling

Umbrella Sampling method was proposed by Torrie and Valleau in 1977,[14] and is still widely used nowadays. Suppose we are studying a transition process between two states such as conversion between two dominant conformations or a chemical reaction, and these two states are separated by a high barrier relative to k_BT. Therefore, the transition is a rare event. A schematic representation of the free energy landscape is shown in Fig. 2.2.

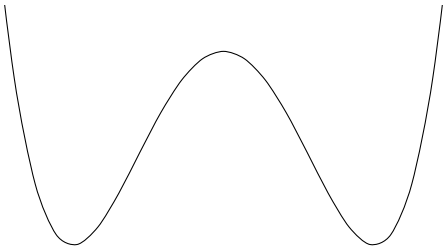

Figure 2.2: A typical free energy surface. Two free energy wells are separated by a barrier higher than k_BT.

Sometimes, we are interested in not only these two dominant states but also the whole pathway. Usually, we define a reaction coordinate ξ and calculate the potential of mean force along this reaction coordinate from the "reactant" to the "product". The reaction coordinate can be either real coordinates such as the difference of bond lengths in, for example, an S_N2 reaction, or a thermodynamics coupling parameter (λ) that defines an unphysical path. However, if we run a simulation with the reaction coordinate initially set to the transition state or the slope, the system will quickly roll back to the "reactant" or the "product" state in order to reduce the free energy. The consequence is that phase space outside the "reactant" and "product" regions cannot be sampled sufficiently to yield accurate free energy profile in a brute force simulation. In order to enhance the exploration in these regions, a series of artificial biasing potentials as (usually harmonic) functions of ξ can be added to the potential energy. Simulations are performed on these potential energy surfaces

$$U_i(\mathbf{R}) = U_0(\mathbf{R}) + \Delta U_i(\xi). \tag{2.3.0.1}$$

Each biased simulation is called a *window*. The strengths of the biases should be strong enough to maintain the system in the vicinity of where you are interested in, and also should be weak enough that the system can have significant overlap in two adjacent windows. After all the simulations, the (biased) distribution of the samples in the whole region should be as flat as possible. Ensemble average under U_0 can be calculated from the ensembles

generated under the biased Hamiltonians U via

$$\begin{aligned}\langle X(\mathbf{R})\rangle_0 =& \frac{\int X(\mathbf{R}) \exp\left[-\beta U_0(\mathbf{R})\right] \mathrm{d}\mathbf{R}}{\int \exp\left[-\beta U_0(\mathbf{R})\right] \mathrm{d}\mathbf{R}} \\ =& \frac{\int X(\mathbf{R}) \exp\left[\beta \Delta U_i(\mathbf{R})\right] \exp\left[-\beta U_i(\mathbf{R})\right] \mathrm{d}\mathbf{R}}{\int \exp\left[\beta \Delta U_i(\mathbf{R})\right] \exp\left[-\beta U_i(\mathbf{R})\right] \mathrm{d}\mathbf{R}} \\ =& \frac{\langle X \exp\left(\beta \Delta U_i\right)\rangle_i}{\langle \exp\left(\beta \Delta U_i\right)\rangle_i}. \end{aligned} \tag{2.3.0.2}$$

Better postprocessing methods are the Weighted Histogram Analysis Method and the Multistate Bennett Acceptance Ratio method (to be discussed in Section 3.1.4 and 3.1.5).

2.4 Adaptive Biasing Force Method

If the conditional gradient of the free energy with respect to a reaction coordinate (mean force) over the equilibrium distribution of the system *restricted* to the hypersurface where the reaction coordinate is constant can be computed, the free energy profile along this specific reaction coordinate can be readily obtained by thermodynamic integration. In the following, we shall follow the derivation by Ciccotti et al.[15] For a system under molecular constraints, $\sigma_j(x) = 0$, $j = 1, \ldots, M$, the probability density reads

$$\rho(x) = Z_\sigma^{-1} e^{-\beta V(x)} \prod_{j=1}^{M} \delta(\sigma_j(x)), \tag{2.4.0.1}$$

in which

$$Z_\sigma = \int e^{-\beta V(x)} \prod_{j=1}^{M} \delta(\sigma_j(x)) \, \mathrm{d}x \tag{2.4.0.2}$$

is the configuration integral. By definition, the free energy associated with the vectorial reaction coordinate $q(x) = (q_1(x), \ldots, q_N(x))$ is given by

$$F(z) \coloneqq -\beta^{-1} \ln \left[Z_\sigma^{-1} \int e^{-\beta V(x)} \prod_{k=1}^{M} \delta(q_k(x) - z_k) \prod_{j=1}^{M} \delta(\sigma_j(x)) \, \mathrm{d}x \right], \tag{2.4.0.3}$$

where $z = (z_1, \ldots, z_N)$. By differentiating both sides with respect to z_j, we find

$$\frac{\partial F(z)}{\partial z_j} = -\beta^{-1} e^{\beta F(z)} \cdot Z_\sigma^{-1} \int e^{-\beta V(x)} \frac{\partial}{\partial z_j} \prod_{k=1}^{N} \delta(q_k(x) - z_k) \times \prod_{j=1}^{M} \delta(\sigma_j(x)) \, \mathrm{d}x. \tag{2.4.0.4}$$

Please note that z_j is a number to which the reaction coordinate is to be constrained. Therefore, $V(x)$ is not a function of z_j.

Notice that

$$\begin{aligned}
&\frac{\partial}{\partial z_j} \prod_{k=1}^{N} \delta(q_k(x) - z_k) \times \prod_{j=1}^{M} \delta(\sigma_j(x)) \\
&= -\delta'(q_j(x) - z_j) \prod_{k \neq j} \delta(q_k(x) - z_k) \times \prod_{j=1}^{M} \delta(\sigma_j(x)) \\
&= -\left(b_j(x) \cdot \nabla \delta(q_j(x) - z_j)\right) \prod_{k \neq j} \delta(q_k(x) - z_k) \times \prod_{j=1}^{M} \delta(\sigma_j(x)) \\
&= -b_j(x) \cdot \nabla \left(\prod_{k=1}^{N} \delta(q_k(x) - z_k) \times \prod_{j=1}^{M} \delta(\sigma_j(x)) \right)
\end{aligned} \tag{2.4.0.5}$$

where $b_j(x), j = 1, \ldots, N$ are vector fields satisfying

$$b_j(x) \cdot \nabla \sigma_k(x) = 0, \quad \forall j = 1, \ldots, N,\ k = 1, \ldots, M \tag{2.4.0.6}$$

and

$$b_j(x) \cdot \nabla q_k(x) = \begin{cases} 1 & \text{if } j = k \\ 0 & otherwise \end{cases}. \tag{2.4.0.7}$$

Thereby,

$$\begin{aligned} &\frac{\partial F(z)}{\partial z_j} \\ &= -\beta^{-1} e^{\beta F(z)} \cdot Z_\sigma^{-1} \int e^{-\beta V(x)} b_j(x) \nabla \left(\prod_{k=1}^{N} \delta(q_k(x) - z_k) \prod_{j=1}^{M} \delta(\sigma_j(x)) \right) \mathrm{d}x \\ &= e^{\beta F(z)} \cdot Z_\sigma^{-1} \int e^{-\beta V(x)} \Big(b_j(x) \cdot \nabla V(x) - \beta^{-1} \nabla \cdot b_j(x) \Big) \\ &\times \prod_{k=1}^{N} \delta(q_k(x) - z_k) \prod_{j=1}^{M} \delta(\sigma_j(x)) \, \mathrm{d}x \end{aligned} \tag{2.4.0.8}$$

after integration by parts. After rearrangement, the gradient of $F(z)$ (i.e. the mean force) can be expressed as

$$\frac{\partial F}{\partial z_j} = \Big\langle b_j(x) \cdot \nabla V - \beta^{-1} \nabla \cdot b_j(x) \Big\rangle_{q(x)=z, \sigma(x)=0}, \tag{2.4.0.9}$$

where $\langle \cdot \rangle_{q(x)=z,\sigma(x)=0}$ denotes the conditional average under the constraints $q(x) = z$, $\sigma(x) = 0$. For any function $f(x)$

$$\langle f \rangle_{q(x)=z,\sigma(x)=0} = \frac{\int f(x) e^{-\beta V(x)} \prod_{k=1}^{M} \delta(q_k(x) - z_k) \prod_{j=1}^{M} \delta(\sigma_j(x)) \, \mathrm{d}x}{\int e^{-\beta V(x)} \prod_{k=1}^{M} \delta(q_k(x) - z_k) \prod_{j=1}^{M} \delta(\sigma_j(x)) \, \mathrm{d}x}. \tag{2.4.0.10}$$

Being "restricted" here is different from being "constrained". In the latter, there is an additional condition that the velocity of this reaction coordinate must be set to zero. In the standard Blue Moon sampling method developed by Carter et al.[16], constrained molecular dynamics is utilized to compute the conditional expectation in Eq. 2.4.0.10. However, it introduces additional constraints on the momenta, which has to be removed. Therefore, computing the mean force from a constrained ensemble, a correction factor (denoted as $|Z|^{-1/2}$ in Ref. [16]) must be introduced, which arises from performing the momentum integration in the ensemble average. In addition, constrained simulation may cause quasinonergodic effect, in particular when multiple reaction pathways are present. Therefore, constrained simulation is not recommended.

Alternatively, adaptive biasing force (ABF) method, which was proposed by Darve and Pohorille in 2001[17] and reformulated in 2008[18], can be

used for the calculations of free energy profiles. It applies to unconstrained simulations, as well as constrained simulations. In ABF, an external force, $-\left\langle F_\xi|_{\xi^*}\right\rangle \nabla\xi$, that counteracts the mean force is applied. The net result of this procedure is that, after a brief equilibrium, the average force acting on ξ is close to zero and the system undergoes barrierless diffusionlike motion along the order parameter. This means that the sampling of ξ becomes uniform.

We denote by $\boldsymbol{\xi}$ the vector of all order parameters ξ_i, $i = 1, \ldots, N_\xi$. The free energy $A(\boldsymbol{\xi})$ is defined as

$$A(\boldsymbol{\xi}) = -\ln \int e^{-H(\mathbf{x})} \prod_{j=1}^{N_\xi} \delta(\xi_j - \xi_j(\mathbf{x}))\, \mathrm{d}\mathbf{x}. \tag{2.4.0.11}$$

β has been absorbed. Now define a thin matrix $\mathbf{W}$ with N_ξ columns, which satisfies

$$\mathbf{J}_{\boldsymbol{\xi}} \mathbf{W} = \mathbf{I}, \tag{2.4.0.12}$$

where $\mathbf{J}_{\boldsymbol{\xi}}$ is a fat matrix with its element defined by

$$[\mathbf{J}_{\boldsymbol{\xi}}]_{ij} = \frac{\partial \xi_i}{\partial x_j}, \tag{2.4.0.13}$$

and $\mathbf{I}$ is a unit matrix. Using the definition of ensemble average and integration by parts, we find

$$\begin{aligned}
\left\langle \mathbf{W}^t \nabla U - (\nabla \cdot \mathbf{W})^t \Big|_{\boldsymbol{\xi}} \right\rangle =& \frac{\int \left(\mathbf{W}^t \nabla U - (\nabla \cdot \mathbf{W})^t\right) e^{-U} \prod_{j=1}^{N_\xi} \delta(\xi_j - \xi_j(\mathbf{x}))\, \mathrm{d}\mathbf{x}}{\int e^{-U(\mathbf{x})} \prod_{j=1}^{N_\xi} \delta(\xi_j - \xi_j(\mathbf{x}))\, \mathrm{d}\mathbf{x}} \\
=& \frac{-\int (\nabla \cdot (e^{-U}\mathbf{W}))^t \prod_{j=1}^{N_\xi} \delta(\xi_j - \xi_j(\mathbf{x}))\, \mathrm{d}\mathbf{x}}{\int e^{-U(\mathbf{x})} \prod_{j=1}^{N_\xi} \delta(\xi_j - \xi_j(\mathbf{x}))\, \mathrm{d}\mathbf{x}} \\
=& \frac{\int e^{-U} \mathbf{W}^t \nabla \left(\prod_{j=1}^{N_\xi} \delta(\xi_j - \xi_j(\mathbf{x}))\right) \mathrm{d}\mathbf{x}}{\int e^{-U(\mathbf{x})} \prod_{j=1}^{N_\xi} \delta(\xi_j - \xi_j(\mathbf{x}))\, \mathrm{d}\mathbf{x}}.
\end{aligned} \tag{2.4.0.14}$$

Let us choose an index i ($1 \le i \le N_\xi$) and focus on $\partial A/\partial \xi_i$. Only row i of $\mathbf{W}^t$, w_i, needs to be considered. The gradient can be computed as

$$\nabla \left(\prod_{j=1}^{N_\xi} \delta(\xi_j - \xi_j(\mathbf{x})) \right) = \sum_{k=1}^{N_\xi} \delta'(\xi_k(\mathbf{x}) - \xi_k) \prod_{j \neq k} \delta(\xi_j - \xi_j(\mathbf{x})) \nabla \xi_k. \tag{2.4.0.15}$$

Since we have $\nabla \xi_k w_i = \delta_{ik}$,

$$w_i \cdot \nabla \left(\prod_{j=1}^{N_\xi} \delta(\xi_j - \xi_j(\mathbf{x})) \right) = \delta'(\xi_i(\mathbf{x}) - \xi_i) \prod_{j \neq i} \delta(\xi_j - \xi_j(\mathbf{x})). \tag{2.4.0.16}$$

Therefore, the ith component of $\left\langle \mathbf{W}^t\nabla U - (\nabla\cdot\mathbf{W})^t\big|_{\boldsymbol{\xi}}\right\rangle$ is

$$\frac{\int e^{-U}\delta'(\xi_i(\mathbf{x})-\xi_i)\prod_{j\neq i}\delta(\xi_j-\xi_j(\mathbf{x}))\,\mathrm{d}\mathbf{x}}{\int e^{-U(\mathbf{x})}\prod_{j=1}^{N_\xi}\delta(\xi_j-\xi_j(\mathbf{x}))\,\mathrm{d}\mathbf{x}} = \frac{\partial A}{\partial\xi_i}, \tag{2.4.0.17}$$

where a property of δ function

$$\int f(x)\delta'(x)\,\mathrm{d}x = -\int f'(x)\delta(x)\,\mathrm{d}x \tag{2.4.0.18}$$

has been used. This proves

$$\nabla_{\boldsymbol{\xi}} A = \left\langle \mathbf{W}^t\nabla U - (\nabla\cdot\mathbf{W})^t\Big|_{\boldsymbol{\xi}}\right\rangle. \tag{2.4.0.19}$$

This can be used in conjunction with the calculations of first and second spatial derivatives.

For multiple reaction coordinates, the calculation of $\nabla_\xi A$ can requires only first derivatives by observing that, with $\mathbf{J}(\mathbf{w})_{ij} = \frac{\partial w_i}{\partial x_j}$,

$$\begin{aligned}\left\langle \frac{\mathrm{d}}{\mathrm{d}t}(w_i\cdot\mathbf{p})\Big|_{\boldsymbol{\xi}}\right\rangle &= \left\langle \mathbf{p}^t\mathbf{M}^{-1}\mathbf{J}(w_i)^t\mathbf{p} - w_i\cdot\nabla U\Big|_{\boldsymbol{\xi}}\right\rangle\\ &= \left\langle -w_i\cdot\nabla U + \mathrm{Tr}(\mathbf{J}(w_i))|_{\boldsymbol{\xi}}\right\rangle\\ &= -\left\langle w_i\cdot\nabla U - \nabla\cdot w_i|_{\boldsymbol{\xi}}\right\rangle\\ &= -\frac{\partial A}{\partial\xi_i}.\end{aligned} \tag{2.4.0.20}$$

During the deviation, the equality

$$\int \mathbf{u}^t\mathbf{B}\mathbf{u}e^{-\mathbf{u}^t\mathbf{A}\mathbf{u}}\,\mathrm{d}\mathbf{u} = \frac{1}{2}\,\mathrm{Tr}\,(A^{-1}B)\int e^{-\mathbf{u}^t\mathbf{A}\mathbf{u}}\,\mathrm{d}\mathbf{u} \tag{2.4.0.21}$$

has been used with $\mathbf{u}=\mathbf{p}$, $\mathbf{B}=\mathbf{M}^{-1}\mathbf{J}(\mathbf{W})^t$, and $\mathbf{A}=\mathbf{M}^{-1}$.

For the choice $\mathbf{W}^t = \mathbf{M}_\xi\mathbf{J}_\xi\mathbf{M}^{-1}$, $\mathbf{M}_\xi^{-1} = \mathbf{J}_\xi\mathbf{M}^{-1}\mathbf{J}_\xi^t$, we get

$$\nabla_{\boldsymbol{\xi}} A = -\left\langle \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{M}_{\boldsymbol{\xi}}\frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}t}\right)\Big|_{\boldsymbol{\xi}}\right\rangle. \tag{2.4.0.22}$$

This equation is much easier to implement numerically than Eq. 2.4.0.19. No second derivatives are involved. This is especially convenient since computing terms like $\partial\mathbf{M}_{\boldsymbol{\xi}}/\partial x_t$ can be quite tedious to implement.

2.5 λ-dynamics

The coupling parameter λ is treated as a pseudo particle with fictitious mass m_λ. The extended Hamiltonian for the system with a coupling parameter in one dimension can be written as

$$H(\mathbf{R}, \lambda) = H_{Rxn}(\mathbf{R}, \lambda) + \frac{m_\lambda}{2}\dot{\lambda}^2 + U^*(\lambda), \tag{2.5.0.1}$$

where H_{Rxn} is a legitimate mapping provided that $H_{Rxn}(\mathbf{R}, \lambda = 0)$ and $H_{Rxn}(\mathbf{R}, \lambda = 1)$ correspond to the Hamiltonians for the reactant and product states respectively, and $U^*(\lambda)$ is a restraint that limits the range of λ. Extension of this method to multiple coupling parameters $\{\lambda_i\}$ is straightforward. The pseudo particles can be coupled to high temperature baths, so it can have strengthened ability to overcome the barrier. However, this might lead to energy transfer between the pseudo degrees of freedom to the configuration degrees of freedom. Therefore, the fictitious mass m_λ should be large enough to make this degree of freedom nearly adiabatic from the rest of the system.[19] λ-dynamics can also be coupled with metadynamics,[20] which will be introduced in Sec. 2.7.

2.6 Wang-Landau Algorithm

Wang-Landau algorithm was developed by Wang and Landau in 2001 to accelerate the convergence in calculating the density of states.[21] In conventional Monte Carlo simulation at a certain temperature T, the configurations are generated with a probability proportional to the product of the density of states $g(E)$ and the Boltzmann factor e^{-E/k_BT}. While Wang-Landau algorithm aims to estimate the density of states $g(E)$ via a random walk in energy space to produce a flat histogram. If a random walk in energy space is performed with a probability proportional to the reciprocal of the density of states $1/g(E)$, a flat histogram can be generated for the energy distribution. This is accomplished by simultaneously modifying the estimated density of states in a systematic way to produce a flat histogram over the allowed range of energy and making the density of states converge to the true value. Note that at the beginning of the random walk, the density of state is normally unknown, so we simply set them to one for all the energies, i.e. $g(E) = 1$. Then the random walk in energy space begins by changing the configuration, for instance flipping the spin in Ising model, randomly with a probability

$$p(E_1 \to E_2) = min\left[\frac{g(E_1)}{g(E_2)}, 1\right], \tag{2.6.0.1}$$

where E_1 and E_2 are energies of the configurations before and after the change. Each time an energy level E is visited, the corresponding density of states is updated by multiplying the existing value by a modification factor $f > 1$, i.e., $g(E) \to g(E)f$. Initially, the modification factor f can be set to a value as large as $f_0 = e^1$, which leads to a crazy exploration in the energy space and the walker can quickly cover all energy levels. This random walk keeps on until we have a "flat" histogram $H(E)$. At this moment, the energy levels have been swept in a coarse manner and the density of states converges to the true value with an accuracy proportional to $\ln(f)$. Now, we reduce the modification factor to a finer one according to some recipe such as $f_1 = \sqrt{f_0}$ (any function that monotonically decrease to 1 will do) and reset the histogram $H(E) = 0$. Then we begin the next round of random walk with a finer modification factor f_1 until the histogram is flat again. This iteration continues with $f_{i+1} = \sqrt{f_i}$ until a prechosen criterion such as $f_{final} < \exp(10^{-8}) = 1.00000001$ has been reached. In reality, a perfect "flatness" can never be reached. But we can define a "flat" histogram to be the condition that the histograms for all the E level is not less than 80% of the average histogram $\langle H(E)\rangle$.

This method can be further enhanced by performing, for instance, multiple random walks etc.

This algorithm was extended by Atchadé and Liu[22], and by Liang et al[23].

2.7 Metadynamics

Metadynamics, vividly called flooding method, was first suggested by Laio and Parrinello in 2002.[24] Imaging you were standing in a valley and were surrounded by high mountains. In most of the time, you were just wandering near the minimum, because your kinetic energy was not enough to climb the mountains. Suddenly, you realized that you could use metadynamics as a magic to escape from the minimum. You started walking. After each step, you took a bottle of sand out of your miraculous pocket and put the sand under your feet. Then you were lifted up inch-by-inch, and the deposited sand piles discourage you from revisiting where you had visited. And you were finally raised up to the top of the mountain and at that moment you were able to climb over that mountain without much effort and fell into another valley. The magic of sand continued, and at last you smoothed the whole area. Because you kept recording where you had put the sand and how much sand you had put there. You drew the shape the piled sand according to the record and you flipped it. In this way, you got the exact shape of the original free energy landscape up to a constant.

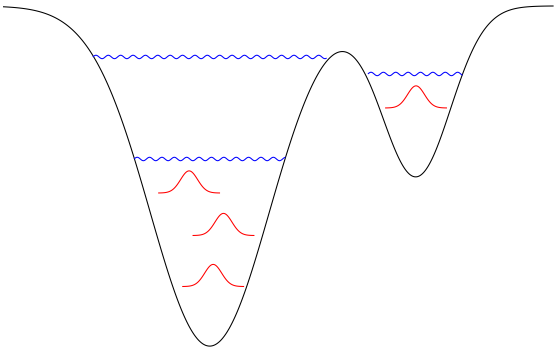

Figure 2.3: A schematic representation of metadynamics. The free energy well is gradually filled up with small Gaussians, and a transition is facilitated.

The above texts are merely an informal explanation of metadynamics. Formally, metadynamics belongs to a class of methods in which sampling is facilitated by introducing additional bias potential to pre-selected degrees of freedom, which are often referred as collective variables (CVs). In metadynamics, the bias potential added to the Hamiltonian of the system is history-dependent, and is often written as a sum of Gaussians deposited during the simulation as

$$V_G(\mathbf{s},t) = \int_0^t \mathrm{d}t' \omega \exp\left(-\sum_{i=1}^{d} \frac{[\mathbf{s}_i(R) - \mathbf{s}_i(R(t'))]^2}{2\sigma_i^2}\right) \tag{2.7.0.1}$$

on a collective variable $\mathbf{s}$ in d-dimension. σ and ω are two parameters tuning the shape of the Gaussians, which can be time-dependent. Asymptotically,

$$V_G(\mathbf{s}, t \to \infty) = -G(\mathbf{s}) + C. \tag{2.7.0.2}$$

Recent improvement over the ordinary metadynamics on convergence issue, which is termed well-tempered metadynamics (WTMetaD), can be found in Ref. [25] and is reviewed in Ref. [26]. In this method, a gradually tempered Gaussian hill

$$V_n(\mathbf{s}) = V_{n-1}(\mathbf{s}) + G(\mathbf{s}, \mathbf{s}_n) \exp\left[-\frac{1}{\gamma - 1}\beta V_{n-1}(\mathbf{s}_n)\right] \tag{2.7.0.3}$$

is accumulated, where $V_0(\mathbf{s}) = 0$ and

$$G(\mathbf{s}, \mathbf{s}') = W \exp\left(-||\mathbf{s} - \mathbf{s}'||^2\right) \tag{2.7.0.4}$$

with $||\mathbf{s} - \mathbf{s}'||^2$ being a distance metric such as

$$||\mathbf{s} - \mathbf{s}'||^2 = \frac{1}{2}\sum_{i,j}(\mathbf{s}_i - \mathbf{s}'_i)\Sigma^{-1}_{i,j}(\mathbf{s}_j - \mathbf{s}'_j). \tag{2.7.0.5}$$

$\Sigma^{-1}_{i,j}$ is the inverse of the covariance matrix $\Sigma_{i,j}$, and the latter is normally diagonal $\Sigma_{i,j} = \delta_{i,j}\sigma_i^2$. W is the height of the Gaussian. $G(\mathbf{s}, \mathbf{s}_n)$ is the biasing kernel centered on the current CV value $\mathbf{s}_n$ and is scaled by $\exp\left[-\frac{1}{\gamma-1}\beta V_{n-1}(\mathbf{s}_n)\right]$ when being accumulated. The scaling factor itself decreases as $1/n$, therefore the change of the biasing potential becomes smaller as the metadynamics simulation progresses.[27]

Practically, the update of the biasing potential is performed every N_G steps. Between any two adjacent updates, the system evolves under the action of the biasing potential $V_n(\mathbf{s}(\mathbf{R}))$. After the nth update, the biasing potential is

$$V(\mathbf{s}, t) = \sum_{k=1}^{n} W \exp\left(-||\mathbf{s} - \mathbf{s}_k||^2\right) \exp\left[-\frac{1}{\gamma - 1}\beta V_{k-1}(\mathbf{s}_k)\right]. \tag{2.7.0.6}$$

The factor $(\gamma - 1)\beta^{-1}$ is sometimes referred to as $k_B\Delta T$.

The remarkable feature of this stochastic update of the biasing potential is that the evolution of the bias can be described asymptotically by an ordinary differential equation (ODE), which after some manipulations reads

$$\frac{\mathrm{d}V(\mathbf{s}, t)}{\mathrm{d}t} = \int \mathrm{d}\mathbf{s}'\, G(\mathbf{s}, \mathbf{s}') \exp\left[-\frac{1}{\gamma - 1}\beta V(\mathbf{s}', t)\right] P_V(\mathbf{s}', t), \tag{2.7.0.7}$$

where

$$P_V(\mathbf{s}, t) = \frac{e^{-\beta[F(\mathbf{s}) + V(\mathbf{s}, t)]}}{\int \mathrm{d}\mathbf{s}'\, e^{-\beta[F(\mathbf{s}') + V(\mathbf{s}', t)]}}. \tag{2.7.0.8}$$

For any $G(\mathbf{s}, \mathbf{s}')$, this ODE has the asymptotic solution

$$V(\mathbf{s}, t) = -\left(1 - \frac{1}{\gamma}\right) F(\mathbf{s}) + c(t), \tag{2.7.0.9}$$

where

$$c(t) = \frac{1}{\beta} \log \frac{\int \mathrm{d}\mathbf{s}\, e^{-\beta F(\mathbf{s})}}{\int \mathrm{d}\mathbf{s}\, e^{-\beta[F(\mathbf{s})+V(\mathbf{s},t)]}} \tag{2.7.0.10}$$

is independent of $\mathbf{s}$. Metadynamics thus converges to the desired result. Two interesting consequences arise.

First, it can be shown that a time-dependent estimator for $F(\mathbf{s})$ is given by

$$F(\mathbf{s}) = -\left(\frac{\gamma}{\gamma - 1}\right) V(\mathbf{s}, t) + \frac{1}{\beta} \log \int \mathrm{d}\mathbf{s}\, \exp\left[\frac{\gamma}{\gamma - 1} \beta V(\mathbf{s}, t)\right]. \tag{2.7.0.11}$$

Taking this equation into Eq. 2.7.0.10, one obtains

$$c(t) = \frac{1}{\beta} \log \frac{\int \mathrm{d}\mathbf{s} \exp\left[\frac{\gamma}{\gamma-1} \beta V(\mathbf{s}, t)\right]}{\int \mathrm{d}\mathbf{s} \exp\left[\frac{1}{\gamma-1} \beta V(\mathbf{s}, t)\right]}. \tag{2.7.0.12}$$

Second, it offers a practical way of calculating the expectation value of any $\mathbf{R}$-dependent function $O(\mathbf{R})$ as the simulation proceeds. The idea is that at time t the biased probability distribution for $\mathbf{R}$ is given by

$$P_V(\mathbf{R}, t) = \frac{e^{-\beta[U(\mathbf{R})+V(\mathbf{s}(\mathbf{R}),t)]}}{\int \mathrm{d}\mathbf{R}\, e^{-\beta[U(\mathbf{R})+V(\mathbf{s}(\mathbf{R}),t)]}}, \tag{2.7.0.13}$$

which can be rewritten as

$$P_V(\mathbf{R}, t) = P(\mathbf{R}) e^{-\beta[V(\mathbf{s}(R),t)-c(t)]}, \tag{2.7.0.14}$$

where $P(\mathbf{R})$ is the unbiased Boltzmann distribution, and $e^{\beta[V(\mathbf{s}(R),t)-c(t)]}$ is the time-dependent unbiasing factor. Straightforwardly, the average of $O(\mathbf{R})$ over the unbiased ensemble can be calculated from the metadynamics trajectory as

$$\langle O(\mathbf{R}) \rangle = \left\langle O(\mathbf{R}) e^{\beta[V(\mathbf{s}(\mathbf{R},t)-c(t))]} \right\rangle_V. \tag{2.7.0.15}$$

This reweighting can be used to obtain the FES for some set of CVs $\mathbf{s}'$ either biased or unbiased by setting $O(\mathbf{R}) = \delta[\mathbf{s}' - \mathbf{s}'(\mathbf{R})]$. It is also useful if one chooses $\mathbf{s}'$ as the biased degree of freedom $\mathbf{s}$ and obtain the FES. Disagreement between the FESs obtained directly from the bias potential and through reweighting is a clear sign that the metadynamics simulation has not converged.

Metadynamics has been implemented in PLUMED (`https://plumed.github.io/doc-v2.3/user-doc/html/_metadyn.html`), which can work with major molecular dynamics packages.

2.8 Variationally Enhanced Sampling Method

Variationally Enhanced Sampling (VES) method was developed by Valsson and Parrinello in 2014 an evolution of Metadynamic.[28] It begins with the following functional of a bias potential $V(\mathbf{s})$

$$\Omega[V] = \frac{1}{\beta} \ln \frac{\int d\mathbf{s} e^{-\beta[F(\mathbf{s})+V(\mathbf{s})]}}{\int d\mathbf{s} e^{-\beta F(\mathbf{s})}} + \int d\mathbf{s} p(\mathbf{s}) V(\mathbf{s}), \tag{2.8.0.1}$$

where $p(\mathbf{s})$ is an arbitrary normalized probability distribution, and $F(\mathbf{s})$ is the unbiased free energy surface. This functional is convex and invariant under the addition of an arbitrary constant to $V(\mathbf{s})$, $\Omega[V + k] = \Omega[V]$. The potential that renders $\Omega[V]$ stationary is, with an constant,

$$V(\mathbf{s}) = -F(\mathbf{s}) - (1/\beta) \ln p(\mathbf{s}) \tag{2.8.0.2}$$

for $p(\mathbf{s}) \neq 0$ and $V(\mathbf{s}) = \infty$ otherwise. This stationary point is also the global minimum of $\Omega[V]$ since the functional is convex.

To make use of the variational property of $\Omega[V]$, the bias potential $V(\mathbf{s})$ is expanded as a function of a set of variational parameters $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_K)$, and then the function $\Omega(\boldsymbol{\alpha}) = \Omega[V(\boldsymbol{\alpha})]$ is minimized with respect to $\boldsymbol{\alpha}$ until convergence is reached. With the converged potential $V(\mathbf{s}; \boldsymbol{\alpha})$, the free energy surface $F(\mathbf{s})$ can be estimated from Eq. 2.8.0.2.

The gradient $\Omega(\boldsymbol{\alpha})'$

$$\frac{\partial \Omega(\boldsymbol{\alpha})}{\partial \alpha_i} = -\left\langle \frac{\partial V(\mathbf{s};\boldsymbol{\alpha})}{\partial \alpha_i} \right\rangle_{V(\boldsymbol{\alpha})} + \left\langle \frac{\partial V(\mathbf{s};\boldsymbol{\alpha})}{\partial \alpha_i} \right\rangle_{p} \tag{2.8.0.3}$$

and the Hessian $\Omega''(\boldsymbol{\alpha})$

$$\begin{aligned} \frac{\partial^2 \Omega(\boldsymbol{\alpha})}{\partial \alpha_j \partial \alpha_i} =& \beta \operatorname{Cov}\left[\frac{\partial V(\mathbf{s};\boldsymbol{\alpha})}{\partial \alpha_j}, \frac{\partial V(\mathbf{s};\boldsymbol{\alpha})}{\partial \alpha_i} \right]_{V(\boldsymbol{\alpha})} \\ & - \left\langle \frac{\partial^2 V(\mathbf{s};\boldsymbol{\alpha})}{\partial \alpha_j \partial \alpha_i} \right\rangle_{V(\boldsymbol{\alpha})} + \left\langle \frac{\partial^2 V(\mathbf{s};\boldsymbol{\alpha})}{\partial \alpha_j \partial \alpha_i} \right\rangle_{p} \end{aligned} \tag{2.8.0.4}$$

where $\langle \cdots \rangle_{V(\boldsymbol{\alpha})}$ and $\operatorname{Cov}[\cdots]_{V(\boldsymbol{\alpha})}$ are the expectation value and the covariance, respectively, obtained in a biased simulation employing the potential $V(\mathbf{s}; \boldsymbol{\alpha})$, and $\langle \cdots \rangle_p$ is an expectation value in the distribution $p(\mathbf{s})$. A natural approach is to expand $V(\mathbf{s}; \boldsymbol{\alpha})$ in a linear basis set and use the coefficient of this expansion as variational parameters,

$$V(\mathbf{s}; \boldsymbol{\alpha}) = \sum_k \alpha_k G_k(\mathbf{s}). \tag{2.8.0.5}$$

In this case the gradient and the Hessian simplify,

$$\frac{\partial \Omega(\boldsymbol{\alpha})}{\partial \alpha_{\mathrm{i}}} = -\left\langle G_i(\mathbf{s}) \right\rangle_{V(\boldsymbol{\alpha})} + \left\langle G_i(\mathbf{s}) \right\rangle_{p}, \tag{2.8.0.6}$$

$$\frac{\partial^2 \Omega(\boldsymbol{\alpha})}{\partial \alpha_j \partial \alpha_i} = \beta \operatorname{Cov}\left[G_j(\mathbf{s}), G_i(\mathbf{s}) \right]_{V(\boldsymbol{\alpha})}. \tag{2.8.0.7}$$

2.9 Orthogonal Space Random Walk

The orthogonal space random walk (OSRW) was developed by Yang in 2008.[29] Phase space sampling is always hindered by free energy barriers. As shown above, several methods have been proposed to accelerate the transition between two states separated by a large free energy barrier, via alchemical process or enhanced conformational switching. In alchemical process, we define a coupling parameter λ. Similarly, in conformational switching we define a reaction coordinate $\mathbf{S}$. Essentially, these two methods are the same, because the coupling parameter λ can be regarded as a coordinate for extended dynamics. Without loss of generality, we can write the free energy difference with the order parameter $\xi = \xi_i$ and $\xi = \xi_f$ as

$$\Delta G(\xi_i \to \xi_f) = \int_{\xi_i}^{\xi_f} \frac{\partial G}{\partial \xi}\bigg|_{\xi'} \mathrm{d}\xi' = \int_{\xi_i}^{\xi_f} \left\langle \frac{\partial U}{\partial \xi} - \beta^{-1} \frac{\partial \ln |J|}{\partial \xi} \right\rangle_{\xi'} \mathrm{d}\xi', \quad (2.9.0.1)$$

where J is the Jacobian term corresponding to the coordinate transformation between the Cartesian coordinates and the reaction coordinates, and $\frac{\partial U}{\partial \xi} - RT\frac{\partial \ln |J|}{\partial \xi}$ can be regarded as the generalized force F_ξ on ξ. Because the transformation from $\xi = \xi_i$ to $\xi = \xi_f$ is slow, we can either constrain or restrain the system to a series of ξ'. Unfortunately, albeit the acceleration along the reaction coordinate, the relaxation of the other degrees of freedom is usually hindered by some "hidden barriers" and is not able to catch up with the alternation of the reaction coordinate. This is called "Hamiltonian lagging" as identified by Kollman et al.[30] Therefore, acceleration of the space orthogonal to the reaction coordinate is equally important as the acceleration of the reaction coordinate.

Orthogonal space random walk is one of the approaches that can deal with this difficulty. In this method, all the coordinates perpendicular to the reaction coordinate are grouped together into F_ξ. A small two dimensional biasing potential $G(\xi, F_\xi)$, instead of a one-dimensional one as in metadynamics (see Sec. 2.7), is added to the Hamiltonian of the system recursively, which has a functional form like

$$h \exp\left(-\frac{|\xi - \xi(t_i)|^2}{2{w_1}^2}\right) \exp\left(-\frac{|F_\xi - F_\xi(t_i)|^2}{2{w_2}^2}\right). \quad (2.9.0.2)$$

The overall biasing potential

$$G(\xi, F_\xi) = \sum_{t_i} h \exp\left(-\frac{|\xi - \xi(t_i)|^2}{2{w_1}^2}\right) \exp\left(-\frac{|F_\xi - F_\xi(t_i)|^2}{2{w_2}^2}\right). \quad (2.9.0.3)$$

will eventually flatten the underlying free energy surface along the orthogonal space. Application of this biasing potential to conformational free energy calculations is straightforward, while for alchemical free energy calculations

it can be realized by λ-dynamics developed by Charlie Brooks.[31] Similar to metadynamics, the free energy profile along the two-dimensional reaction coordinates $[\xi(t_i), F_\xi]$ can be estimated as $-G\left(\xi, F_\xi\right) + C$, where C is an irrelevant constant. Correspondingly, the generalized force distribution at ξ' should be proportional to $\exp\left[\beta G\left(\xi', F_{\xi'}\right)\right]$, and the free-energy derivative can be obtained via

$$\left.\frac{\partial G}{\partial \xi}\right|_{\xi'} = \langle F_\xi \rangle_{\xi'} = \frac{\sum F_\xi \exp\left[\beta G(\xi, F_\xi)\right]\delta(\xi - \xi')}{\sum \exp\left[\beta G(\xi, F_\xi)\right]\delta(\xi - \xi')}, \tag{2.9.0.4}$$

which can be fed into the thermodynamic integration formula to obtain the free energy change from $\xi = \xi_i$ to any target state with the order parameter ξ as the following

$$\Delta G(\xi) = \int_{\xi_i}^{\xi} \left.\frac{\partial G}{\partial \xi}\right|_{\xi'} \mathrm{d}\xi'. \tag{2.9.0.5}$$

2.10 Enveloping Distribution Sampling

Enveloping distribution sampling method was first proposed by Christ and van Gunsteren in 2007.[32] When calculating the free energy difference between states A and B,

$$\Delta G_{BA} = G_B - G_A = -\beta^{-1} \ln \frac{Q_B}{Q_A}, \tag{2.10.0.1}$$

we may encounter convergence difficulty if the important spaces of these two states are well separated, shown as black lines in Fig. 2.4. Simulation under the Hamiltonian of state A can hardly cover the important region of Hamiltonian B, and then the free energy of state B will be significantly overestimated.

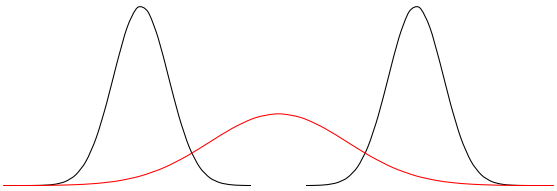

Figure 2.4: The configuration distributions under two Hamiltonians have no visible overlap as shown by solid black curves. A reference state (shown as the red curve) that has remarkable overlap with both states can be introduced to accelerate the convergence of the free energy calculations using, for instance, TP.

A simple solution to this difficulty is "overlap sampling", in which a reference state that can cover the important regions of both Hamiltonians A and B is introduced. We then carry out a simulation for the reference state and the free energy difference between state A and B can be calculated as

$$\Delta G_{BA} = \Delta G_{BR} - \Delta G_{AR} = -\beta^{-1} \ln \frac{\langle e^{-\beta(H_B - H_R)} \rangle_R}{\langle e^{-\beta(H_A - H_R)} \rangle_R}, \tag{2.10.0.2}$$

which is a combination of two thermodynamic perturbation calculations from the reference state to the target states.

However, building the Hamiltonian of the reference state is not trivial. Without knowledge of the Hamiltonians for state A and state B, we cannot generate an effective Hamiltonian, especially in a high dimensional space. Enveloping distribution sampling method provides a natural way to generate the Hamiltonian for the reference state with simply mixing the Hamiltonians of state A and state B in the following way

$$H_R(\mathbf{r}) = -(s\beta)^{-1} \ln \left(e^{-s\beta H_A(\mathbf{r})} + e^{-s\beta H_B(\mathbf{r})} \right), \tag{2.10.0.3}$$

where s is a scale factor that modulates the mixing[33] as shown in Fig. 2.5. Increasing s lowers the barrier height separating the two minima in the mixed potential, thereby enhances the transition. Straightforwardly, you may come to the idea that running Hamiltonian-REMD with different s can remarkably increase the efficiency. If you take a close look at Eq. 2.10.0.3, you will find that s appears always with β. In other words, changing s is equivalent to changing the temperature for the simulation. This is one interesting case where H-REMD and T-REMD are coincident with each other.

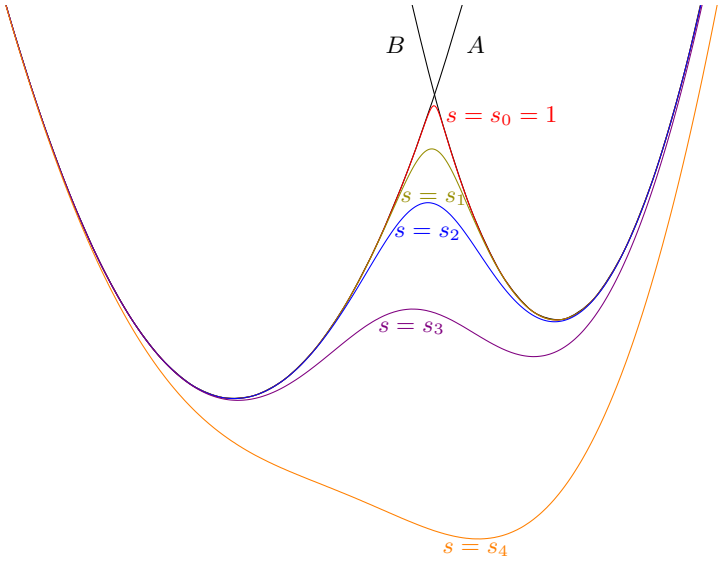

Figure 2.5: State A and state B have only negligible overlap at high energy regions. The reference state generated by the mixing of state A and state B is characterized by s. Increasing s may lower the barrier between the dominant wells.

The force is also a mixing quantity from two Hamiltonians as

$$\mathbf{F}_R^i = -\frac{\partial H_R}{\partial \mathbf{r}^i} = \frac{e^{-s\beta H_A(\mathbf{r})}}{e^{-s\beta H_A(\mathbf{r})} + e^{-s\beta H_B(\mathbf{r})}} \left(-\frac{\partial H_A(\mathbf{r})}{\partial \mathbf{r}^i}\right) + \frac{e^{-s\beta H_B(\mathbf{r})}}{e^{-s\beta H_A(\mathbf{r})} + e^{-s\beta H_B(\mathbf{r})}} \left(-\frac{\partial H_B(\mathbf{r})}{\partial \mathbf{r}^i}\right). \quad (2.10.0.4)$$

2.11 String Method

2.11.1 Zero Temperature String Method

The zero temperature string method was developed by E, Ren and Vanden-Eijnden in 2002.[34] In 2007, they simplified this method by eliminating the projecting the potential force onto the hyperplane perpendicular to the string.[35] The overall algorithm is an iterative application of a simple two-step procedure: (I) evolution of the string by standard ordinary differential equation (ODE) solvers, and (II) the reparametrization of the string by interpolation. The accuracy of this method is determined by the 2nd step, while its efficiency is determined by the 1st step.

The minimum energy path (MEP) is the most probable path that the system will take under the overdamped dynamics to move between two minima at a and b on the potential energy surface $V(x)$, crossing the barriers in-between. Let's denote the MEP by γ, which satisfies

$$(\nabla V)^{\perp}(\gamma) = 0, \tag{2.11.1.1}$$

where $(\nabla V)^{\perp}$ is the component of (∇V) normal to γ,

$$(\nabla V)^{\perp}(\gamma) = \nabla V(\gamma) - (\nabla V(\gamma), \hat{\tau})\,\hat{\tau}. \tag{2.11.1.2}$$

Here $\hat{\tau}$ denotes the unit tangent of the curve γ, and $(\cdot,\cdot)$ denotes the Euclidean inner product.

The string method locates the MEP by evolving a curve connecting a and b, under the potential force field. The simplest dynamics for the evolution is given abstractly by

$$\nu_n = -(\nabla V)^{\perp}, \tag{2.11.1.3}$$

where ν_n denotes the normal velocity of the curve. Only the normal component of the velocity matters for the evolution of a curve, while the tangential velocity only moves points along the curve, changing the parametrization of the curve without changing the curve itself.

First, we take a particular parametrization of the curve $\gamma : \gamma = \{\boldsymbol{\varphi}(\alpha) : \alpha \in [0,1]\}$. Then we have $\hat{\tau}(\alpha) = \boldsymbol{\varphi}_\alpha/|\boldsymbol{\varphi}_\alpha|$, where $\boldsymbol{\varphi}_\alpha$ denotes the derivative of $\boldsymbol{\varphi}$ with respect to α. The simplest parametrization is equal arc-length parametrization, in which α is a constant multiple of the arc length from a to the point $\boldsymbol{\varphi}(\alpha)$. In this case, we also have $|\boldsymbol{\varphi}_\alpha| = \text{const}$ (this constant being the length of the curve γ).

The string method evolves the curve via

$$\dot{\boldsymbol{\varphi}} = \nabla V(\boldsymbol{\varphi}) + \bar{\lambda}\hat{\tau}, \tag{2.11.1.4}$$

where $\bar{\lambda}(\alpha,t)\hat{\tau}(\alpha,t)$ is a Lagrange multiplier term for the purpose of enforcing the particular parametrization of the string. The string is discretized

into a number of images $\{\boldsymbol{\varphi}_i(t), i = 0, 1, \ldots, N\}$. The images along the string are evolved by iterating upon the following two-step procedure based on time splitting of the terms at the right hand side of Eq. 2.11.1.4.

In the first step, the discrete point on the string are evolved over some time interval Δt according to the full potential force,

$$\dot{\boldsymbol{\varphi}}_i = -\nabla V(\boldsymbol{\varphi}_i). \tag{2.11.1.5}$$

This equation can be integrated in time by any suitable ODE solver. If we denote the positions of the images after n iterations of the scheme by $\boldsymbol{\varphi}_i^n$, $i = 0, \ldots, N$, the new set of images are given by

$$\boldsymbol{\varphi}_i^* = \boldsymbol{\varphi}_i^n - \Delta t \nabla V(\boldsymbol{\varphi}_i^n). \tag{2.11.1.6}$$

In the second step, this new set of images are redistributed along the string using a simple interpolation/reparametrization procedure. Two possible schemes for reparametrization can be applied.

Parametrization by equal arc length Given the values $\{\boldsymbol{\varphi}_i^*\}$ on a nonuniform mesh $\{\alpha_i^*\}$, these values are interpolated onto a uniform mesh with the same number of points via two steps:

1. The arc length corresponding to the current images,

$$s_0 = 0, \quad s_i = s_{i-1} + |\boldsymbol{\varphi}_i^* - \boldsymbol{\varphi}_{i-1}^*|, \quad i = 1, 2, \ldots, N. \tag{2.11.1.7}$$

   The mesh $\{\alpha_i^*\}$ is then obtained by normalizing $\{s_i\}$,

$$\alpha_i^* = s_i / s_N. \tag{2.11.1.8}$$

2. The points $\boldsymbol{\varphi}_i^{n+1}$ at the uniform grids points $\alpha_i = i/N$ are obtained by interpolation. This can be done, for example, by using cubic spline interpolation for the data $\{(\alpha_i^*, \boldsymbol{\varphi}_i^*), i = 0, \ldots, N\}$

Parametrization by energy-weighted arc length The energy-weighted arc length parameterization gives finer resolution around the saddle points, and thus better estimate of the energy barrier and also the unstable direction at those points than the equal arc length scheme does. In this scheme, the energy-weighted arc length corresponding to the current images are computed,

$$s_0^w = 0, \quad s_i^w = s_{i-1}^w + W_{i-(1/2)}|\boldsymbol{\varphi}_i^* - \boldsymbol{\varphi}_{i-1}^*|, \quad i = 1, 2, \ldots, N. \tag{2.11.1.9}$$

Here $W_{i-(1/2)} = W(V_{i+1/2})$ and $V_{i+1/2}$ is the average of the potential energy at $\boldsymbol{\varphi}_{i-1}^*$ and $\boldsymbol{\varphi}_i^*$. The weight function $W(z)$ is some positive, increasing function of $z \in \mathbb{R}$. The mesh $\{\alpha_i^*\}$ is obtained by normalizing $\{s_i^w\} : \alpha_i^* = s_i^w / s_N^w$. The new points $\boldsymbol{\varphi}_i$ on $\alpha_i = i/N$ are then calculated by cubic spline interpolation across the points $\{\boldsymbol{\varphi}_i^*, i = 0, \ldots, N\}$.

Once the new points $\{\boldsymbol{\varphi}_i^{n+1}, i = 0, \ldots, N\}$ are calculated, the algorithm goes back to step 1 and iterates until convergence.

2.11.2 Finite Temperature String Method

Finite temperature string method is an extension of the zero temperature string method. Let $\boldsymbol{\varphi}(\alpha)$ be a curve in configuration space parametrized by $\alpha \in [0,1]$ whose end points, $\boldsymbol{\varphi}(0)$ and $\boldsymbol{\varphi}(1)$, belong to the two metastable sets. In order to have $\boldsymbol{\varphi}(\alpha)$ converges towards the center of the effective reaction tube, an ensemble of realizations $\{\boldsymbol{\varphi}^{\omega}(\alpha)\}$ are introduced and their mean is defined to be the string, i.e., $\langle \boldsymbol{\varphi}^{\omega}(\alpha) \rangle \equiv \boldsymbol{\varphi}(\alpha)$. Each realization evolves by

$$\boldsymbol{\varphi}_t^{\omega} = -\left(\nabla V(\boldsymbol{\varphi}^{\omega})\right)^{\perp} + (\eta^{\omega})^{\perp} + r\hat{t}. \tag{2.11.2.1}$$

Here $\hat{t} = \boldsymbol{\varphi}_{\alpha}/|\boldsymbol{\varphi}_{\alpha}|$ is the unit tangent vector along $\boldsymbol{\varphi}$ and $a^{\perp} = a - (\hat{t} \cdot a)\hat{t}$ is the projection of the vector a in the hyperplane normal to the string $\boldsymbol{\varphi}(\alpha)$ denoted here by $S(\alpha)$. η^{ω} is a white noise which covariance

$$\left\langle \eta_j^{\omega}(\alpha,t)\eta_k^{\omega}(\alpha',0) \right\rangle = \begin{cases} 2k_B T\delta_{jk}\delta t & \text{if } \alpha = \alpha' \\ 0 & \text{otherwise} \end{cases} \tag{2.11.2.2}$$

The scalar field $r \equiv r(\alpha,t)$ is a Lagrange multiplier term to preserve some particular parametrization of the string $\boldsymbol{\varphi}$ chosen beforehand, for instance the equal arc length or equal energy-weighted arc length.

The equilibrium density function for Eq. 2.11.2.1 is given by

$$\rho(\mathbf{q},\alpha) = Z^{-1}(\alpha)e^{-\beta V(\mathbf{q})}\delta_{S(\alpha)}(\mathbf{q}) \tag{2.11.2.3}$$

where $\delta_{S(\alpha)}(\mathbf{q})$ is the Dirac distribution concentrated on $S(\alpha)$, and $Z(\alpha) = \int_{S(\alpha)} e^{-\beta V(\mathbf{q})}\,\mathrm{d}\sigma$ is the normalization constant. By definition, the center of the transition tube is given by

$$\boldsymbol{\varphi}(\alpha) = Z^{-1}(\alpha)\int_{S(\alpha)} \mathbf{q}e^{-\beta V(\mathbf{q})}\,\mathrm{d}\sigma. \tag{2.11.2.4}$$

The width of the effective transition tube itself can be characterized by a few times the variance of $\mathbf{q}$ around $\boldsymbol{\varphi}(\alpha)$; i.e., its local radius square can be defined as

$$R^2(\alpha) = \lambda Z^{-1}(\alpha)\int_{S(\alpha)} |\mathbf{q} - \boldsymbol{\varphi}(\alpha)|^2 e^{-\beta V(\mathbf{q})}\,\mathrm{d}\sigma, \tag{2.11.2.5}$$

where λ is a number of order unity. The integral of $\rho(\mathbf{q},\alpha)$ in the ball of radius $R(\alpha)$ centered around $\rho(\alpha)$ gives the probability that a dynamical trajectory involved in the transition event crosses the plane $S(\alpha)$ within this ball.

The free energy is given by

$$F(\alpha) = -k_B T \ln \int_{S(\alpha)} e^{-\beta V(\mathbf{q})}\,\mathrm{d}\sigma. \tag{2.11.2.6}$$

Using thermodynamic integration, the free energy difference between α_1 and α_2 becomes

$$F(\alpha_2) - F(\alpha_1) = \int_{\alpha_1}^{\alpha_2} \left\langle (\hat{t} \cdot \nabla V)((\hat{t} \cdot \boldsymbol{\varphi})_\alpha - \hat{t} \cdot \boldsymbol{\varphi}) \right\rangle \mathrm{d}\alpha. \tag{2.11.2.7}$$

In this implementation, constrained molecular dynamics must be carried out for the calculations of ensemble averages over S_α, which is usually difficult. In 2009, Vanden-Eijnden and Venturoli updated this FTS algorithm by replacing the hyperplanes perpendicular to the string with Voronoi cells.[36] This new algorithm begins with an initial set of images, $\boldsymbol{\varphi}_\alpha^0$ with $\alpha = 0, \ldots, N$ with equal arc length, i.e. $|\boldsymbol{\varphi}_{\alpha+1}^0 - \boldsymbol{\varphi}_\alpha^0| = |\boldsymbol{\varphi}_\alpha^0 - \boldsymbol{\varphi}_{\alpha-1}^0|$ for all $\alpha = 1, \ldots, N-1$. Each image is associated with a replica of the original system, $\mathbf{x}_\alpha^0$, $(\boldsymbol{\varphi}(\mathbf{x}_\alpha^0) = \boldsymbol{\varphi}_\alpha^0)$. Then the positions of these systems are updated iteratively for $n \geq 0$ upon the following steps:

1. Update $\mathbf{x}_\alpha^n$ with a reflecting boundary condition at the boundary of the Voronio cell associated with the image $\boldsymbol{\varphi}_\alpha^n$ via, for example,

$$\mathbf{x}_\alpha^* = \mathbf{x}_\alpha^n - \Delta t \nabla V(\mathbf{x}_\alpha^n) + \sqrt{2\beta^{-1}\Delta t}\boldsymbol{\xi}_\alpha^n \tag{2.11.2.8}$$

   and set

$$\mathbf{x}_\alpha^{n+1} = \begin{cases} \mathbf{x}_\alpha^*, & \text{if } \mathbf{x}_\alpha^* \in B_\alpha^n \\ \mathbf{x}_\alpha^n, & \text{otherwise} \end{cases} \tag{2.11.2.9}$$

   where

$$B_\alpha^n = \{\mathbf{x} \text{ such that } |\boldsymbol{\varphi}(\mathbf{x}) - \boldsymbol{\varphi}_\alpha^n| < |\boldsymbol{\varphi}(\mathbf{x}) - \boldsymbol{\varphi}_{\alpha'}^n| \text{ for all } \alpha' \neq \alpha\} \tag{2.11.2.10}$$

   Δt denotes the time step used for numerical integration and $\boldsymbol{\xi}_\alpha^n$ are independent Gaussian variables with mean 0 and variance 1.

2. Compute the running average of each $\mathbf{x}_\alpha^n$,

$$\bar{\mathbf{x}}_\alpha^n = \frac{1}{n} \sum_{n'=0}^{n} \mathbf{x}_\alpha^{n'}. \tag{2.11.2.11}$$

3. Update each image along the string toward the running average $\bar{\mathbf{x}}_\alpha^n$ while keeping the string smooth. To do so use

$$\boldsymbol{\varphi}_\alpha^* = \boldsymbol{\varphi}_\alpha^n - \Delta\tau(\boldsymbol{\varphi}_\alpha^n - \boldsymbol{\varphi}(\bar{\mathbf{x}}_\alpha^n)) + \mathbf{r}_\alpha^*, \tag{2.11.2.12}$$

   where $\Delta\tau > 0$, $\mathbf{r}_0^* = \mathbf{r}_N^* = 0$, and for $\alpha = 1, \ldots, N-1$,

$$\mathbf{r}_\alpha^* = k^n(\boldsymbol{\varphi}_{\alpha+1}^* + \boldsymbol{\varphi}_{\alpha-1}^* - 2\boldsymbol{\varphi}_\alpha^*). \tag{2.11.2.13}$$

   Here $k^n > 0$ is an adjustable parameter, whose value controls how aggressive the smoothing is.

4. Enforcing the equal arc-length parametrization by interpolating a piecewise linear curve through $\{\boldsymbol{\varphi}^*_\alpha\}_{\alpha=0,\dots,N}$ and redistributing points at equal distance along this curve to obtain $\{\boldsymbol{\varphi}^{n+1}_\alpha\}_{\alpha=0,\dots,N}$.

5. If $\mathbf{x}^{n+1}_\alpha \in B^{n+1}_\alpha$ go to step 1, otherwise set $\mathbf{x}^{n+1}_\alpha = \boldsymbol{\varphi}^{n+1}_\alpha$ and go to step 1. Repeat until convergence of $\{\boldsymbol{\varphi}^{n+1}_\alpha\}_{\alpha=0,\dots,N}$.

3

Postprocessing

"The source of mistake is always between the keyboard and the chair. So, check, double check and check again."

– Gerhard König

3.1 Rigorous Methods

3.1.1 Thermodynamic Perturbation

Thermodynamic Perturbation (TP), also known as Free Energy Perturbation (FEP), exponential average, or Zwanzig equation, was developed by Zwanzig,[37] and Landau and Lifshitz, independently, and probably by Peierls[38].

A reference system containing N-particles can be described by Hamiltonian $H_0(\mathbf{x}, \mathbf{p}_x)$, which is a function of $3N$ Cartesian coordinates, $\mathbf{x}$, and their conjugated momenta, $\mathbf{p}_x$. Similarly, the target system can be described by Hamiltonian $H_1(\mathbf{x}, \mathbf{p}_x)$. These two systems are related by

$$H_1(\mathbf{x}, \mathbf{p}_x) = H_0(\mathbf{x}, \mathbf{p}_x) + \Delta H(\mathbf{x}, \mathbf{p}_x) \tag{3.1.1.1}$$

The Helmholtz free energy difference between the target and the reference systems, ΔA, can be given in terms of the ratio of the corresponding partition functions, Q_1 and Q_0:

$$\Delta A = -\frac{1}{\beta} \ln \frac{Q_1}{Q_0}, \tag{3.1.1.2}$$

where $\beta = (k_B T)^{-1}$, and

$$Q_i = \frac{1}{h^{3N} N!} \iint \exp\left[-\beta H_i(\mathbf{x}, \mathbf{p}_x)\right] \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_\mathbf{x}. \tag{3.1.1.3}$$

Taking Eq. 3.1.1.3 into Eq. 3.1.1.2, we obtain

$$\Delta A = -\frac{1}{\beta}\ln\frac{\iint \exp\left[-\beta H_1(\mathbf{x},\mathbf{p}_x)\right]\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_\mathbf{x}}{\iint \exp\left[-\beta H_0(\mathbf{x},\mathbf{p}_x)\right]\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_\mathbf{x}} \tag{3.1.1.4}$$

$$= -\frac{1}{\beta}\ln\frac{\iint \exp\left[-\beta\Delta H(\mathbf{x},\mathbf{p}_x)\right]\exp\left[-\beta H_0(\mathbf{x},\mathbf{p}_x)\right]\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_\mathbf{x}}{\iint \exp\left[-\beta H_0(\mathbf{x},\mathbf{p}_x)\right]\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_\mathbf{x}}, \tag{3.1.1.5}$$

The probability density function of finding the reference system in a state defined by positions $\mathbf{x}$ and momenta $\mathbf{p}_x$ is

$$P_0(\mathbf{x},\mathbf{p}_x) = \frac{\exp[-\beta H_0(\mathbf{x},\mathbf{p}_x)]}{\iint \exp[-\beta H_0(\mathbf{x},\mathbf{p}_x)]\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_\mathbf{x}} \tag{3.1.1.6}$$

If the probability density function is used, Eq. 3.1.1.5 becomes

$$\Delta A = -\frac{1}{\beta}\ln\iint \exp[-\beta\Delta H(\mathbf{x},\mathbf{p}_x)]P_0(\mathbf{x},\mathbf{p}_x)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{p}_\mathbf{x}, \tag{3.1.1.7}$$

or, equivalently,

$$\Delta A = -\frac{1}{\beta}\ln\left\langle\exp\left[-\beta\Delta H(\mathbf{x},\mathbf{p}_x)\right]\right\rangle_0, \tag{3.1.1.8}$$

Here, $\langle\cdots\rangle_0$ denotes an ensemble average over configurations sampled from the reference state. Equation 3.1.1.8 is the basic equation of TP. It states that ΔA can be estimated by sampling only equilibrium configurations of the reference state.

Note that integration over the kinetic term in the partition function, Eq. 3.1.1.3, can be carried out analytically. Thus, it cancels out in Eq. 3.1.1.2, and Eq. 3.1.1.8 becomes

$$\Delta A_f = -\frac{1}{\beta}\ln\left\langle\exp(-\beta\Delta U)\right\rangle_0, \tag{3.1.1.9}$$

where ΔU is the difference in the potential energy between the target and the reference states. The subscript f is an indication of a forward ($0 \to 1$) TP calculation. The integration implied by the statistical average is now carried out over particle coordinates only. The variance of ΔA is

$$\delta^2\Delta A_f = \frac{1}{N_0\beta^2}\left(\frac{\left\langle(\exp(-\beta\Delta U))^2\right\rangle_0}{\left(\left\langle\exp(-\beta\Delta U)\right\rangle_0\right)^2} - 1\right). \tag{3.1.1.10}$$

If we exchange the reference and the target systems, and repeat the same derivation, using the same convention for ΔA and ΔU as before, we have a backward TP expression for the free energy difference

$$\Delta A_b = \frac{1}{\beta}\ln\left\langle\exp(\beta\Delta U)\right\rangle_1, \tag{3.1.1.11}$$

and the variance is

$$\delta^2 \Delta A_b = \frac{1}{N_1 \beta^2} \left(\frac{\langle (\exp(\beta \Delta U))^2 \rangle_1}{(\langle \exp(\beta \Delta U) \rangle_1)^2} - 1 \right). \tag{3.1.1.12}$$

Although expressions Eq. 3.1.1.9 and Eq. 3.1.1.11 are formally equivalent, their convergence properties may be quite different. This means that there is a preferred direction to carry out the required transformation between the two states. One should start the perturbation from the state having larger important region in phase space. In other words, the reference system should be the one with higher entropy, and the transformation should proceed in the direction in which the entropy decreases. If we have the free energy differences from both the forward and backward TP calculations, we can compute the "best estimate" of ΔA as

$$\Delta A = \frac{(\delta^2 \Delta A_f)^{-1} \Delta A_f + (\delta^2 \Delta A_b)^{-1} \Delta A_b}{(\delta^2 \Delta A_f)^{-1} + (\delta^2 \Delta A_b)^{-1}}, \tag{3.1.1.13}$$

with variance

$$\delta^2 \Delta A = \frac{1}{(\delta^2 \Delta A_f)^{-1} + (\delta^2 \Delta A_b)^{-1}}. \tag{3.1.1.14}$$

Since ΔA is calculated as the average over a quantity that depends only on ΔU, this average can be computed over probability distribution $P_0(\Delta U)$ instead of $P_0(\mathbf{x}, \mathbf{p}_x)$. Then, ΔA in Eq. 3.1.1.7 can be expressed as a one-dimensional integral over energy difference

$$\Delta A = -\frac{1}{\beta} \ln \int \exp(-\beta \Delta U) P_0(\Delta U) \, \mathrm{d}\Delta U, \tag{3.1.1.15}$$

If U_0 and U_1 were functions of a sufficient number of identically distributed random variable, ΔU would follow a Gaussian distribution, which is a consequence of the central limit theorem. In practice, the probability distribution $P_0(\Delta U)$ deviates somewhat from an ideal Gaussian case, but still has a "Gaussian-like" shape. This indicates that the value of the integral in Eq. 3.1.1.15 depends on the low-energy tail of the distribution. Using the language of Jarzynski for the nonequilibrium work[39], the maximum of $P_0(\Delta U)$ is the typical energy difference, and the peak value of $P_0(\Delta U) \cdot \exp(-\beta \Delta U)$ is the dominant realization that contributes the most to ΔA. It clearly shows in Fig. 3.1 that the dominant realization lies to the left of the typical one.

Even though $P_0(\Delta U)$ is only rarely an exact Gaussian, it is instructive to consider this case in more detail. If we substitute

$$P_0(\Delta U) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\Delta U - \langle \Delta U \rangle_0)^2}{2\sigma^2} \right] \tag{3.1.1.16}$$

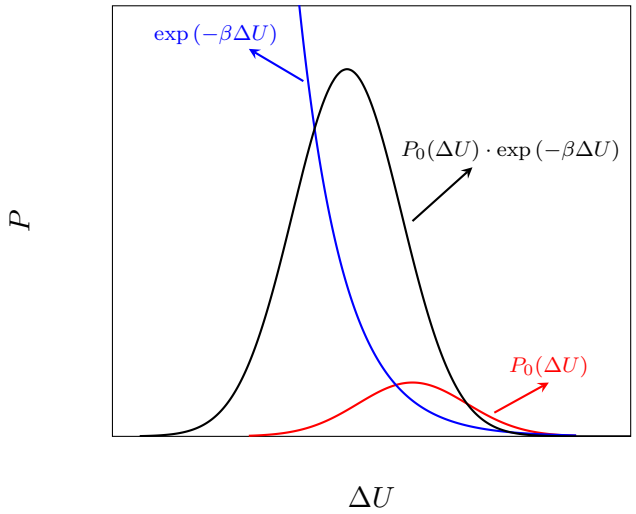

Figure 3.1: $P_0(\Delta U)$, the Boltzmann factor $\exp(-\beta\Delta U)$ and their product, which is the integrand in Eq. 3.1.1.15. The low-ΔU tail of the integrand is poorly sampled with $P_0(\Delta U)$ and, therefore, is known with low statistical accuracy. However, it provides an important contribution to the integral.

where

$$\sigma^2 = \left\langle \Delta U^2 \right\rangle_0 - \langle \Delta U \rangle_0^2 \tag{3.1.1.17}$$

to Eq. 3.1.1.15, we obtain

$$\exp(-\beta\Delta A) = \frac{C}{\sqrt{2\pi}\sigma} \int \exp\left[-\frac{(\Delta U - \langle \Delta U \rangle_0 - \beta\sigma^2)^2}{2\sigma^2} \right] \mathrm{d}\Delta U \tag{3.1.1.18}$$

Here, C is independent of ΔU

$$C = \exp\left[-\beta \left(\langle \Delta U \rangle_0 - \frac{1}{2}\beta\sigma^2 \right) \right] \tag{3.1.1.19}$$

If $P_0(\Delta U)$ is Gaussian, the integral in Eq. 3.1.1.18 can be evaluated analytically using cumulant expansion (see appendix D)

$$\Delta A = \langle \Delta U \rangle_0 - \frac{1}{2}\beta\sigma^2. \tag{3.1.1.20}$$

If the distribution of ΔU deviates from Gaussian, there will be extra terms measuring the skewness of Gaussian. With the leading term, ΔA becomes

$$\Delta A = \langle \Delta U \rangle_0 - \frac{1}{2}\beta\sigma^2 + \frac{\beta^2}{6}\left(\left\langle \Delta U^3 \right\rangle_0 - 3 \left\langle \Delta U^2 \right\rangle_0 \langle \Delta U \rangle_0 + 2 \langle \Delta U \rangle_0^3 \right). \tag{3.1.1.21}$$

3.1.2 Thermodynamic Integration

Thermodynamic Integration (TI) method was proposed by Kirkwood.[40]. If the free energy, A, is a continuous function of λ, the free energy difference between two states corresponding to $\lambda = 0$ and $\lambda = 1$ can be computed via

$$\Delta A = \int_0^1 \frac{\partial A(\lambda)}{\partial \lambda}\, \mathrm{d}\lambda. \tag{3.1.2.1}$$

With

$$A(\lambda) = -\beta^{-1} \ln Q(\lambda), \tag{3.1.2.2}$$

the partial derivative can be expressed as

$$\begin{aligned} \frac{\partial A(\lambda)}{\partial \lambda} =& -\beta^{-1} \left[\frac{\partial \ln Q(\lambda)}{\partial \lambda}\right] \\ =& -\frac{\beta^{-1}}{Q(\lambda)} \frac{\partial Q(\lambda)}{\partial \lambda}. \end{aligned} \tag{3.1.2.3}$$

From the definition of Q

$$Q_{NVT}(\lambda) = \frac{1}{h^{3N} N!} \iint \exp[-\beta H(\mathbf{x}, \mathbf{p}_x, \lambda)]\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_\mathbf{x}, \tag{3.1.2.4}$$

we have

$$\begin{aligned} \frac{\partial Q(\lambda)}{\partial \lambda} =& \frac{1}{h^{3N} N!} \iint \frac{\partial}{\partial \lambda} \exp[-\beta H(\mathbf{x}, \mathbf{p}_x, \lambda)]\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_\mathbf{x} \\ =& -\frac{\beta}{h^{3N} N!} \iint \frac{\partial H(\mathbf{x}, \mathbf{p}_x, \lambda)}{\partial \lambda} \exp[-\beta H(\mathbf{x}, \mathbf{p}_x, \lambda)]\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_\mathbf{x}. \end{aligned} \tag{3.1.2.5}$$

Substituting back into the expression for $\partial A/\partial \lambda$ yields

$$\begin{aligned} \frac{\partial A(\lambda)}{\partial \lambda} =& \frac{1}{h^{3N} N!} \frac{1}{Q(\lambda)} \iint \frac{\partial H(\mathbf{x}, \mathbf{p}_x, \lambda)}{\partial \lambda} \exp[-\beta H(\mathbf{x}, \mathbf{p}_x, \lambda)]\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_\mathbf{x}, \\ =& \frac{1}{h^{3N} N!} \iint \frac{\partial H(\mathbf{x}, \mathbf{p}_x, \lambda)}{\partial \lambda} \cdot \frac{\exp[-\beta H(\mathbf{x}, \mathbf{p}_x, \lambda)]}{Q(\lambda)}\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_\mathbf{x}, \\ =& \left\langle \frac{\partial H(\mathbf{x}, \mathbf{p}_x, \lambda)}{\partial \lambda} \right\rangle_\lambda \end{aligned} \tag{3.1.2.6}$$

Thus, the basic TI formula is

$$\Delta A = \int_0^1 \left\langle \frac{\partial H(\mathbf{x}, \mathbf{p}_x, \lambda)}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda \tag{3.1.2.7}$$

where $\langle \cdots \rangle_\lambda$ corresponds to the ensemble average obtained using the Hamiltonian $H(\lambda)$. In practice, the ensemble of configurations can be obtained by molecular dynamics or Monte Carlo simulations. It is common practice in free energy calculations to use the coupling parameter λ for defining the

transformation from the initial state A with Hamiltonian H_A to the final state B with Hamiltonian H_B. The simplest coupling is linear transformation as

$$H(\lambda) = (1 - \lambda)H_A + \lambda H_B, \tag{3.1.2.8}$$

for which

$$\frac{\partial H(\lambda)}{\partial \lambda} = H_B - H_A. \tag{3.1.2.9}$$

The accuracy of TI integral formula depends on the exactness of the numerical integration method.[41] Practically, the integrand in Eq. 3.1.2.7 needs to be evaluated over a number of discrete points λ_i, and then be summed up to give the free energy difference between $\lambda = 0$ and 1, for instance via the trapezoidal rule

$$\Delta A = \sum_{i=0}^{N-1} \frac{1}{2} \left(\left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda_i} + \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda_{i+1}} \right) (\lambda_{i+1} - \lambda_i). \tag{3.1.2.10}$$

A finite number of λ_i values between 0 and 1 are chosen and for each of them a complete molecular dynamics simulation is carried out resulting in an ensemble of configurations generated with $H(\lambda_i)$. The ensemble average of the derivative of the Hamiltonian with respect to λ is then calculated for each λ_i.

In addition to summation method, the simplest numerical integration is to evaluate the integrand at the midpoint:

$$\Delta A \simeq \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda=\frac{1}{2}} \tag{3.1.2.11}$$

This might be a good first thing to do to get some impression of what is going on, but is only accurate for very smooth or small changes.

3.1.3 Bennett Acceptance Ratio

"The thing that differentiates scientists is purely an artistic ability to discern what is a good idea, what is a beautiful idea, what is worth spending time on, and most importantly, what is a problem that is sufficiently interesting, yet sufficiently difficult, that is hasn't yet been solved, but the time for solving it has come now."
– Professor Savas Dimopoulos, Stanford University

Bennett acceptance ratio was developed by Bennett in 1976,[42] and was re-discovered by Crooks[43] for Markovian and balanced dynamics and by Shirts et al[44] using maximum likelihood over 20 years later. The Metropolis function is defined as

$$M(x) = min\{1, \exp(-x)\}, \tag{3.1.3.1}$$

which has the property

$$M(x)/M(-x) = \exp(-x). \tag{3.1.3.2}$$

If we make a trial move that keeps the same configuration $(q_1, \cdots, q_N)$ but switches the potential function from U_0 to U_1 or vice-versa, the acceptance probabilities for such a pair of trial moves must satisfy the detailed balance

$$M(U_1 - U_0)\exp(-U_0) = M(U_0 - U_1)\exp(-U_1). \tag{3.1.3.3}$$

Integrating this identity over all of configuration space and multiplying by the trivial factors Q_0/Q_0 and Q_1/Q_1, one obtains:

$$Q_0 \frac{\int M(U_1 - U_0)\exp(-U_0)\,\mathrm{d}\mathbf{q}}{Q_0} = Q_1 \frac{\int M(U_0 - U_1)\exp(-U_1)\,\mathrm{d}\mathbf{q}}{Q_1}, \tag{3.1.3.4}$$

or simply

$$\frac{Q_0}{Q_1} = \frac{\langle M(U_0 - U_1)\rangle_1}{\langle M(U_1 - U_0)\rangle_0}. \tag{3.1.3.5}$$

The physical meaning of this formula is that a Monte Carlo calculation that includes potential-switching trial moves would distribute configurations between U_1 and U_0 in the ratio of their configurational integrals.

A formula more general than Eq. 3.1.3.5 can be written as

$$\frac{Q_0}{Q_1} = \frac{Q_0}{Q_1}\frac{\int W\exp(-U_0 - U_1)\,\mathrm{d}\mathbf{q}}{\int W\exp(-U_1 - U_0)\,\mathrm{d}\mathbf{q}} = \frac{\langle W\exp(-U_0)\rangle_1}{\langle W\exp(-U_1)\rangle_0}, \tag{3.1.3.6}$$

where W is an arbitrary weighting function.

Optimization of the free energy estimate is most easily carried out in the limit of large sample sizes. Let the available data consist of n_0 statistically independent configurations from the U_0 ensemble and n_1 from the U_1 ensemble, and let the data be used in Eq. 3.1.3.6 to obtain a finite-sample estimate of the reduced free energy difference $\Delta A = A_1 - A_0 = \ln(Q_0/Q_1)$. Using the error propagation law of uncorrelated variables ($covar(x_1, x_2) = 0$),[45]

$$\delta^2 [y(x_1, x_2)] = \left(\frac{\partial y}{\partial x_1}\right)^2 \delta^2(x_1) + \left(\frac{\partial y}{\partial x_2}\right)^2 \delta^2(x_2). \tag{3.1.3.7}$$

Thus we have the variance of ΔA

$$\begin{aligned} \delta^2(\Delta A) &= \left(\frac{\partial \Delta A}{\partial Q_0}\right)^2 \delta^2 Q_0 + \left(\frac{\partial \Delta A}{\partial Q_1}\right)^2 \delta^2 Q_1 \\ &= \left(\frac{1}{Q_0}\right)^2 \delta^2 Q_0 + \left(-\frac{1}{Q_1}\right)^2 \delta^2 Q_1 \\ &= \left(\frac{1}{Q_0}\right)^2 \delta^2 Q_0 + \left(\frac{1}{Q_1}\right)^2 \delta^2 Q_1. \end{aligned} \tag{3.1.3.8}$$

With the definition of variance $\delta^2 X = \langle X^2 \rangle - \langle X \rangle^2$, we have

$$\begin{aligned} \delta^2 Q_0 &= \delta^2 \langle W \exp(-U_0) \rangle_1 \\ &= \delta^2 \left(\frac{1}{n_1} \sum_{i=1}^{n_1} W_i \exp(-U_0(i)) \right) \\ &= \sum_{i=1}^{n_1} \left(\frac{1}{n_1}\right)^2 \delta^2 (W_i \exp(-U_0(i))) \\ &= \frac{1}{n_1} \delta^2 (W_i \exp(-U_0(i))) \\ &= \frac{1}{n_1} \left\{ \left\langle [W \exp(-U_0)]^2 \right\rangle_1 - [\langle W \exp(-U_0) \rangle_1]^2 \right\} \\ &= \frac{1}{n_1} \left\{ \left\langle W^2 \exp(-2U_0) \right\rangle_1 - [\langle W \exp(-U_0) \rangle_1]^2 \right\}, \end{aligned} \tag{3.1.3.9}$$

which shows that the variance of the mean of the samples equals to the variance of the samples divided by the number of samples.

With sufficiently large sample sizes, the error of this estimate will be nearly Gaussian, and its expected square is exactly the variance of ΔA

$$\begin{aligned} &\delta^2(\Delta A_{est} - \Delta A) \\ \approx &\frac{\langle W^2 \exp(-2U_1) \rangle_0}{n_0 [\langle W \exp(-U_1) \rangle_0]^2} + \frac{\langle W^2 \exp(-2U_0) \rangle_1}{n_1 [\langle W \exp(-U_0) \rangle_1]^2} - \frac{1}{n_0} - \frac{1}{n_1} \\ = &\frac{\int [(Q_0/n_0) \exp(-U_1) + (Q_1/n_1) \exp(-U_0)] W^2 \exp(-U_0 - U_1) \, \mathrm{d}\mathbf{q}}{[\int W \exp(-U_0 - U_1) \, \mathrm{d}\mathbf{q}]^2} \\ &- \frac{1}{n_0} - \frac{1}{n_1}. \end{aligned} \tag{3.1.3.10}$$

To minimize it with respect to W, we have

$$W = const \times \left(\frac{Q_0}{n_0} \exp\left(-U_1\right) + \frac{Q_1}{n_1} \exp\left(-U_0\right) \right)^{-1}. \tag{3.1.3.11}$$

Substituting this into Eq. 3.1.3.6 yields

$$\frac{Q_0}{Q_1} = \frac{\langle f(U_0 - U_1 + C) \rangle_1}{\langle f(U_1 - U_0 - C) \rangle_0} \exp\left(+C\right), \tag{3.1.3.12}$$

where

$$C = \ln \frac{Q_0 n_1}{Q_1 n_0}, \tag{3.1.3.13}$$

and f denotes the Fermi function

$$f(x) = \frac{1}{1 + \exp\left(+x\right)}. \tag{3.1.3.14}$$

It can also be expressed as[44]

$$n_1 \langle f(U_0 - U_1 + C) \rangle_1 = n_0 \langle f(U_1 - U_0 - C) \rangle_0. \tag{3.1.3.15}$$

It should be noted that Eq. 3.1.3.12 is true for any C, which is actually a shift for one of the potential function. But the particular value specified in Eq. 3.1.3.13 minimizes the expected square error given the finite numbers (n_0 and n_1) of samples.

The variance of ΔA can be obtained by substituting Eq. 3.1.3.11 into Eq. 3.1.3.10, and is

$$\begin{aligned} \delta^2 \Delta A =& \frac{\left\langle f^2(U_1 - U_0 - C) \right\rangle_0}{n_0 \left\langle f(U_1 - U_0 - C) \right\rangle_0^2} + \frac{\left\langle f^2(U_0 - U_1 + C) \right\rangle_1}{n_1 \left\langle f(U_0 - U_1 + C) \right\rangle_1^2} - \frac{1}{n_0} - \frac{1}{n_1} \\ =& \left(\int \frac{n_0 n_1 \rho_0 \rho_1}{n_0 \rho_0 + n_1 \rho_1} \, \mathrm{d}\mathbf{q} \right)^{-1} - \frac{n_0 + n_1}{n_0 n_1}, \end{aligned} \tag{3.1.3.16}$$

in which $\rho_i = \exp\left(-U_i\right)/Q_i$ is the probability.

It is worth emphasizing that Bennett acceptance ratio is asymptotically unbiased, and no other asymptotically unbiased estimator has lower asymptotic mean-squared error. However, it is not clear whether its behavior is always better than other estimators with finite sample sizes.

3.1.4 Weighted Histogram Analysis Method

The weighted histogram analysis method is a generalization of the histogram method developed by Ferrenberg and Swendsen.[46]

Weighted Histogram Analysis Method for Parallel Tempering

The following derivation quite follows Ref. [47]. One of the central quantities in statistical mechanics is configurational integral Z, which in textbook is often written as

$$Z = \int \exp\left(-\beta U(\mathbf{R})\right) \mathrm{d}\mathbf{R}. \tag{3.1.4.1}$$

This is an integral in coordinate space. It also can be written as an integral in energy space

$$Z = \int \Omega(U) \exp\left(-\beta U\right) \mathrm{d}U, \tag{3.1.4.2}$$

where $\Omega(U)$ is density of states and $\Omega(U)\Delta U$ is the number of states in the region $U - \Delta U/2 < U < U + \Delta U/2$. Accordingly, the statistical expectation of an operator $\mathbf{A}$ can be calculated by

$$\langle \mathbf{A} \rangle = \frac{\int \mathbf{A}(U)\Omega(U) \exp\left(-\beta U\right) \mathrm{d}U}{\int \Omega(U) \exp\left(-\beta U\right) \mathrm{d}U}, \tag{3.1.4.3}$$

where

$$\mathbf{A}(U') = \frac{\int \delta(U(\mathbf{R}) - U')\mathbf{A}(\mathbf{R}) \mathrm{d}\mathbf{R}}{\int \delta(U(\mathbf{R}) - U') \mathrm{d}\mathbf{R}}, \tag{3.1.4.4}$$

is the average of $\mathbf{A}$ over all the samples with an energy U'. Therefore, the core objective is to calculate $\Omega(U)$.

Suppose we have one trajectory with N snapshots denoted as $\{\mathbf{R}_n\}$. We then discretize the energy space into M bins with width ΔU, and count the number of snapshots into each bin. For convenience, we define $\psi_m(U)$ as

$$\psi_m(U) = \begin{cases} 1 & \text{if } U \in [U_m - \Delta U/2, U_m + \Delta U/2) \\ 0 & \text{otherwise} \end{cases} \tag{3.1.4.5}$$

Then the histogram for the mth energy bin is

$$H_m = \sum_{n=1}^{N} \psi_m(U(\mathbf{R}_n)) = N \cdot \frac{1}{N} \sum_{n=1}^{N} \psi_m(U(\mathbf{R}_n)) = N \cdot \langle \psi_m \rangle, \tag{3.1.4.6}$$

with variances (see Appendix A)

$$\begin{aligned} \delta^2 H_m =& N^2 \delta^2(\langle \psi_m \rangle) \\ =& g_m N \left(\left\langle {\psi_m}^2 \right\rangle - \langle \psi_m \rangle^2 \right) \\ =& g_m N \left(\langle \psi_m \rangle - \langle \psi_m \rangle^2 \right) \\ =& g_m H_m \left(1 - \frac{H_m}{N} \right). \end{aligned} \tag{3.1.4.7}$$

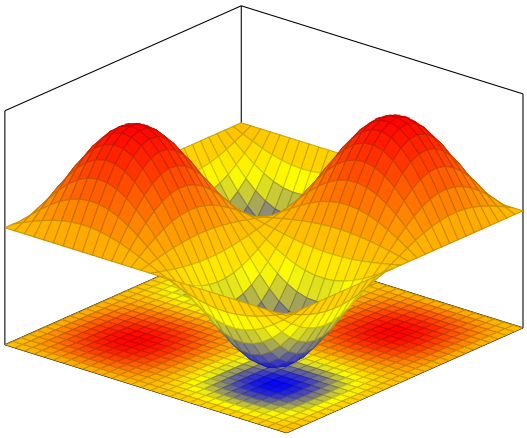

Figure 3.2: A sample histogram in 2D space, for instance potential energy and a reaction coordinate ξ.

A sample histogram in 2D space is shown in Fig. 3.2.

The ratio of the histogram H_m to the total number of snapshots N divided by the bin width ΔU can be approximately taken as the probability of states in this bin, i.e.,

$$\frac{\Omega_m \exp\left(-\beta U_m\right)}{Z} \approx \frac{H_m}{N\Delta U}. \tag{3.1.4.8}$$

Therefore,

$$\begin{aligned}\Omega_m =& \frac{1}{\Delta U} \cdot \frac{H_m}{N} \cdot \frac{Z(\beta)}{\exp\left(-\beta U_m\right)} \\ =& \frac{H_m}{N\Delta U \exp\left[f - \beta U_m\right]},\end{aligned} \tag{3.1.4.9}$$

and variances

$$\delta^2\Omega_m = \frac{\delta^2 H_m}{\left(N\Delta U \exp\left[f - \beta U_m\right]\right)^2}, \tag{3.1.4.10}$$

in which we have defined a dimensionless free energy $f = -\ln Z(\beta)$.

Practically, we may run multiple (K) trajectories using, for example, replica exchange molecular dynamics simulations. For each trajectory (index k), we have unique estimators for the histogram H_{mk}, the density of states Ω_{mk} and their variances $\delta^2 H_{mk}$ and $\delta^2\Omega_{mk}$ being

$$H_{mk} = \sum_{n=1}^{N_k} \psi_m(U(\mathbf{R}_{kn})), \tag{3.1.4.11}$$

$$\delta^2 H_{mk} = g_{mk} H_{mk} \left(1 - \frac{H_{mk}}{N_k}\right), \tag{3.1.4.12}$$

$$\Omega_{mk} = \frac{H_{mk}}{N_k \Delta U \exp\left[f_k - \beta_k U_m\right]}, \tag{3.1.4.13}$$

and

$$\delta^2 \Omega_{mk} = \frac{\delta^2 H_{mk}}{\left(N_k \Delta U \exp\left[f_k - \beta_k U_m\right]\right)^2}, \tag{3.1.4.14}$$

The optimum estimator of the density of states from all the simulations is

$$\Omega_m = \frac{\sum_{k=1}^{K} \left[\delta^2 \Omega_{mk}\right]^{-1} \Omega_{mk}}{\sum_{k=1}^{K} \left[\delta^2 \Omega_{mk}\right]^{-1}}, \tag{3.1.4.15}$$

which is the weighted average of density of states of all the trajectories with the weight reversely proportional to the variances (see Appendix B).

To make the expression simpler, here we take some approximations. First, normally the energy space is split into a large number of bins. The histogram in each bin is much smaller than the total number of snapshots, i.e. $H_{mk} \ll N_k$. With this approximation, we have

$$\delta^2 H_{mk} \approx g_{mk} H_{mk}. \tag{3.1.4.16}$$

The expectation of H_{mk} can be related to the optimum estimator of the density of states, i.e.

$$\overline{H_{mk}} = N_k \Delta U \Omega_m \exp\left(f_k - \beta_k U_m\right). \tag{3.1.4.17}$$

Then we have

$$\delta^2 H_{mk} = g_{mk} N_k \Delta U \Omega_m \exp\left(f_k - \beta_k U_m\right) \tag{3.1.4.18}$$

and

$$\delta^2 \Omega_{mk} = \frac{\Omega_m}{{g_{mk}}^{-1} N_k \Delta U \exp\left(f_k - \beta_k U_m\right)}. \tag{3.1.4.19}$$

Taking Eq. 3.1.4.13 and Eq. 3.1.4.19 into Eq. 3.1.4.15, we find

$$\Omega_m = \frac{\sum_{k=1}^{K} {g_{mk}}^{-1} H_{mk}}{\sum_{k=1}^{K} {g_{mk}}^{-1} N_k \Delta U \exp\left(f_k - \beta_k U_m\right)}, \tag{3.1.4.20}$$

in which

$$f_k = -\ln \int \Omega(U) \exp\left(-\beta_k U\right) \mathrm{d}U = -\ln \sum_{m=1}^{M} \Omega_m \Delta U \exp\left(-\beta_k U_m\right). \tag{3.1.4.21}$$

Obviously, Eq. 3.1.4.20 and Eq. 3.1.4.21 must be solved iteratively. Applying the error propagation rule to Eq. 3.1.4.20 and using Eq. 3.1.4.18, the variance of Ω_m is given by

$$\delta^2\Omega_m = \frac{\Omega_m}{\sum_{k=1}^{K} {g_{mk}}^{-1} N_k \Delta U \exp\left(f_k - \beta_k U_m\right)}. \tag{3.1.4.22}$$

Using the density of states and its variance, we can estimate the expectation of any configuration function $A(\mathbf{R})$ at any inverse temperature β

$$\langle A \rangle_\beta \approx \frac{\sum_{m=1}^{M} \Omega_m \Delta U \exp\left(-\beta U_m\right) A_m}{\sum_{m=1}^{M} \Omega_m \Delta U \exp\left(-\beta U_m\right)}, \tag{3.1.4.23}$$

where

$$A_m = \frac{\int \mathrm{d}\mathbf{R} A(\mathbf{R}) \psi_m(U(\mathbf{R}))}{\int \mathrm{d}\mathbf{R} \psi_m(U(\mathbf{R}))}. \tag{3.1.4.24}$$

Using histograms of bin m from all the simulations and defining $H_m = \sum_{k=1}^{K} H_{mk}$, an estimator of A_m denoted as $\hat{A}_m$ can be calculated as

$$\hat{A}_m = {H_m}^{-1} \sum_{k=1}^{K} \sum_{n=1}^{N_k} \psi_m(U(\mathbf{R}_{kn})) A(\mathbf{R}_{kn}). \tag{3.1.4.25}$$

Taking Eq. 3.1.4.25 into Eq. 3.1.4.23, we obtain an estimator of $\hat{A}(\beta)$

$$\hat{A}(\beta) = \frac{\sum_{m=1}^{M} \Omega_m \Delta U \exp\left(-\beta U_m\right) {H_m}^{-1} \sum_{k=1}^{K} \sum_{n=1}^{N_k} \psi_m(U(\mathbf{R}_{kn})) A(\mathbf{R}_{kn})}{\sum_{m=1}^{M} \Omega_m \Delta U \exp\left(-\beta U_m\right)} \tag{3.1.4.26}$$

$$= \frac{\sum_{m=1}^{M} \Omega_m \Delta U \exp\left(-\beta U_m\right) {H_m}^{-1} \sum_{k=1}^{K} \sum_{n=1}^{N_k} \psi_m(U(\mathbf{R}_{kn})) A(\mathbf{R}_{kn})}{\sum_{m=1}^{M} \Omega_m \Delta U \exp\left(-\beta U_m\right) {H_m}^{-1} \sum_{k=1}^{K} \sum_{n=1}^{N_k} \psi_m(U(\mathbf{R}_{kn}))} \tag{3.1.4.27}$$

$$= \frac{\sum_{k=1}^{K} \sum_{n=1}^{N_k} w_{kn}(\beta) A_{kn}}{\sum_{k=1}^{K} \sum_{n=1}^{N_k} w_{kn}(\beta)}, \tag{3.1.4.28}$$

where the per-configuration weights $w_{kn}(\beta)$ is given by

$$w_{kn}(\beta) = \sum_{m=1}^{M} {H_m}^{-1} \psi_m(U(\mathbf{R}_{kn})) \Omega_m \exp\left(-\beta U_m\right) \tag{3.1.4.29}$$

Weighted Histogram Analysis Method From Minimizing Statistical Error

In this section, the "traditional" derivation method of WHAM are briefly reviewed.[48] In the WHAM, the goal is to get an optimal unbiased probability distribution $\rho_0(\eta)$, where η is a series of discretized histogram bins indexed by $j = 1, 2, 3, ..., M$ along a certain reaction coordinate. WHAM can be used to analyze the Umbrella Sampling (US) simulations, where a set of simulations indexed by i or $k = 1, 2, 3, ..., S$ are performed with a series of biasing potentials added on the reaction coordinate η. To consider a reference molecular system with the potential energy $U_0(\mathbf{x})$, where $\mathbf{x}$ is the set of atomic coordinates. The reaction coordinate η is a function of the atomic coordinates, i.e. $\eta(\mathbf{x})$. To suppose that the ith molecular simulation has been performed using potential energy function

$$U_i^{(b)}(\eta) = U_0(\mathbf{x}) + W_i(\eta(\mathbf{x})), \tag{3.1.4.30}$$

where $W_i(\eta(\mathbf{x}))$ is the biasing potential added on the reaction coordinate η, e.g. $W_i(\eta) = \frac{1}{2}k_i(\eta - \eta_i)^2$ in a harmonic form. From these simulations a set of normalized biased probability distributions $\rho_i^{(b)}(\eta)$ can be obtained.

$$\rho_i^{(b)}(\eta) = \frac{e^{-\beta U_i^{(b)}(\eta)}}{Q_i^{(b)}}, \tag{3.1.4.31}$$

where $Q_i^{(b)} = \int e^{-\beta U_i^{(b)}(\eta)}\, \mathrm{d}\eta = e^{-\beta f_i^{(b)}}$ and $f_i^{(b)}$ is the biased free energy. The corresponding unnormalized unbiased probability distribution $\rho_i^{(u)}(\eta)$ from the ith simulation is defined as,

$$\rho_i^{(u)}(\eta) = e^{\beta[W_i(\eta) - f_i^{(b)}]}\rho_i^{(b)}(\eta) \tag{3.1.4.32}$$

In the following, the free energy $f_i^{(b)}$ is assumed to be known. It has been shown that in the WHAM method, the total normalized unbiased probability distribution $\rho_0(\eta)$ can be obtained by a linear η-dependent combination of the unbiased histograms $\rho_i^{(u)}(\eta)$

$$\rho_0(\eta) = C\sum_{i=1}^{S} p_i(\eta)\rho_i^{(u)}(\eta), \tag{3.1.4.33}$$

where C is the normalization factor. p_i is the weight to be optimized, which is under a constraint that

$$\sum_{i=1}^{S} p_i(\eta) = 1. \tag{3.1.4.34}$$

These weights are chosen so as to minimize the statistical error made on the total unbiased probability distribution $\rho_0(\eta)$, that is, for any given value of

η,

$$\frac{\partial(\sigma^2[\rho_0(\eta)])}{\partial p_i(\eta)} = 0. \tag{3.1.4.35}$$

It can be easily found that $\rho_0(\eta)$ satisfy

$$\begin{aligned}\rho_0(\eta) =& C\sum_{i=1}^{S}\frac{N_i e^{-\beta[W_i(\eta)-f_i^{(b)}]}}{\sum_{k=1}^{S}N_k e^{-\beta[W_k(\eta)-f_k^{(b)}]}}\rho_i^{(u)}(\eta)\\ =& C\sum_{i=1}^{S}\frac{N_i}{\sum_{k=1}^{S}N_k e^{-\beta[W_k(\eta)-f_k^{(b)}]}}\rho_i^{(b)}(\eta)\\ =& C\frac{\sum_{i=1}^{S}N_i\rho_i^{(b)}(\eta)}{\sum_{k=1}^{S}N_k e^{-\beta[W_k(\eta)-f_k^{(b)}]}},\end{aligned} \tag{3.1.4.36}$$

where $\rho_i^{(b)}(\eta)$ can be written as a δ function,

$$\rho_i^{(b)}(\eta) \equiv \frac{1}{N_i}\sum_{l=1}^{N_i}\delta(\eta-\eta_{i,l}), \tag{3.1.4.37}$$

where $\eta_{i,l}$ is the reaction coordinates of the lth configuration in the ith biased simulation .

Until now, the treatment assumes that the free energy parameters $f_i^{(b)}$ are known. In fact, these parameters can be obtained self-consistently. Indeed, the definition of the free energy $f_i^{(b)}$ is,

$$\begin{aligned}e^{-\beta f_i^{(b)}} =& \int e^{-\beta U_i^{(b)}(\eta)}\,\mathrm{d}\eta\\ =& \int \rho_0(\eta)e^{-\beta W_i(\eta)}\,\mathrm{d}\eta\\ =& C\int\frac{\sum_{i=1}^{S}N_i\rho_i^{(b)}(\eta)}{\sum_{k=1}^{S}N_k e^{-\beta[W_k(\eta)-f_k^{(b)}]}}e^{-\beta W_i(\eta)}\,\mathrm{d}\eta\end{aligned} \tag{3.1.4.38}$$

The set of parameters $f_i^{(b)}$ appear on both sides of Eq. 3.1.4.38, which must be solved iteratively with an initial guess of $f_i^{(b)}$ until convergence is reached. The unbiased free energy corresponding to the histogram can be calculated by

$$f_0(\eta) = -\beta^{-1}\ln\rho_0(\eta) \tag{3.1.4.39}$$

with W in Eq. 3.1.4.38 being 0. The constant C in Eq. 3.1.4.36 is irrelevant, which only causes a constant shift to the free energy profiles. To get rid of it, one may subtract the offset constant $f_0(\eta_1)$ from all the $f_0(\eta_j)$.

Weighted Histogram Analysis Method From Maximum Likelihood

The following derivation quite follows Ref. [49], in which maximum likelihood principle is utilized. Suppose we have performed K simulations, each at a different inverse temperature β_k and possibly with different biasing potential $w_k(\mathbf{R})$. We then discretize the 2D plane spanned by the coordinate and unbiased potential energy into bins, each characterized by $\mathbf{R}_j$ and E_h. To make the following derivation cleaner, we map the 2D bins to one dimensional series with index $l, l = 1, \ldots, L$. Next, we construct histograms for bins using all the samples from the simulations. The probability of finding the system in bin l during the kth simulation can be written as

$$p_{k,l} = f_k c_{k,l} p_l^0, \tag{3.1.4.40}$$

in which p_l^0 is the (simulation-independent) unbiased probability,

$$\begin{aligned} c_{k,l} =& \exp\left[-\beta_k \left(E_l + w_{k,l}\right) + \beta_0 E_l\right] \\ =& \exp\left[-\left(\beta_k - \beta_0\right) E_l\right] \exp\left(-\beta_k w_{k,l}\right) \end{aligned} \tag{3.1.4.41}$$

is the bias factor, E_l is the unbiased energy of bin l, $f_k = \left\{\sum_l c_{k,l} p_l^0\right\}^{-1}$ is the normalization factor. If we expand the expressions for $c_{k,l}$ and p_l^0, we find

$$\begin{aligned} f_k \approx& \left(\sum_l \exp\left[-\beta_k \left(E_l + w_{k,l}\right) + \beta_0 E_l\right] \frac{\exp\left(-\beta_0 E_l\right)}{\sum\limits_j \exp\left(-\beta_0 E_j\right)}\right)^{-1} \\ =& \left(\frac{\sum\limits_l \exp\left[-\beta_k (E_l + w_{k,l})\right]}{\sum\limits_j \exp\left(-\beta_0 E_j\right)}\right)^{-1} \\ \approx& \frac{Z_0}{Z_k}, \end{aligned} \tag{3.1.4.42}$$

which is approximately the ratio of two configurational integrals.

It is worth emphasizing that the biasing potential can be multiple dimensional as, for instance, in a two-dimensional umbrella sampling. If the biasing is only in temperature-space as in replica exchange molecular dynamics

$$c_{k,l} = \exp\left[-\left(\beta_k - \beta_0\right) E_l\right], \tag{3.1.4.43}$$

while if the biasing is only in potential energy as in umbrella sampling

$$c_{k,l} = \exp\left(-\beta_0 w_{k,l}\right). \tag{3.1.4.44}$$

If we assume that each count in each histogram is independent, then the likelihood of observing the kth histogram distribution is given by the

multinomial distribution

$$P(n_{k,1}, n_{k,2}, \ldots, n_{k,L}|p_{k,1}, p_{k,2}, \ldots, p_{k,L}) = \frac{\left(\sum_l n_{k,l}\right)!}{\prod_l n_{k,l}!} \prod_{l=1}^{L} (p_{k,l})^{n_{k,l}} \propto \prod_{l=1}^{L} \left(f_k c_{k,l} p_l^0\right)^{n_{k,l}}. \quad (3.1.4.45)$$

For all K simulations, the likelihood is the product of multinomial

$$P(n_{1,1}, \ldots, n_{1,L}; \ldots; n_{K,1}, \ldots, n_{K,L}|p_1^0, \ldots, p_L^0) \propto \prod_{k=1}^{K} \prod_{l=1}^{L} \left(f_k c_{k,l} p_l^0\right)^{n_{k,l}}, \quad (3.1.4.46)$$

where the likelihood is conditional only on the unbiased probabilities p_l^0, since the bias factors $c_{k,l}$ are known parameters, and the normalization constants f_k are known conditional on p_l^0. The maximum likelihood estimate of the unbiased probabilities can be found by maximizing P in Eq. 3.1.4.46 with respect to $p_1^0, \ldots, p_L^0$ and are given by solutions of the simultaneous nonlinear equations

$$p_l^0 = \frac{\sum_{k=1}^{K} n_{k,l}}{\sum_{k=1}^{K} N_k f_k c_{k,l}} \text{ (for each } l) \quad (3.1.4.47)$$

and

$$f_k = \left\{ \sum_l c_{k,l} p_l^0 \right\}^{-1}, \quad (3.1.4.48)$$

where N_k is the total number of counts in the kth histogram.

Binless Weighted Histogram Analysis Method

The following derivation quite follows Ref. [50]. Let us start with the definition of a generalized energy function u and its corresponding coefficient θ. For instance, for canonical ensemble, $u = U(x)$ is the potential energy function, and $\theta = \beta$ is the inverse temperature. For isothermal grand-canonical ensemble, $u = (U(x), N)$, and $\theta = (\beta, \beta\mu)$, in which N is the number of particles and μ is the chemical potential. For temperature replica exchange molecular dynamics, $u = U(x)$ and $\theta_k = \beta_k$ for the kth replica. For umbrella sampling, $u = (U_0(x), \omega_1(x), \omega_2(x), \ldots, \omega_d(x))$, where $U_0(x)$ is the unbiased Hamiltonian and $\omega_k(x)$ is the biasing potential in window k. Correspondingly, $\theta_k = (\beta, 0, \ldots, 0, \beta, 0, \ldots, 0)$, in which all the elements are zero except for the first and the $(k+1)$th element.

Assume that simulations are conducted at m coefficient vectors θ_r ($r = 1, \ldots, m$) and with the same energy vector $u(x)$. (Note that in this notation the dimensionality, d, of the θ and u vectors and the number of simulations, m, are, in general, distinct. For instance, for temperature replica exchange molecular dynamics as shown above, $d = 1$, while m is the number of replicas.) Denoted by $\{x_{ji} : i = 1, \ldots, n_j\}$ the set of configurations of size n_j from the the jth simulation, and denoted by $u_{ji} = u(x_{ji})$ the corresponding generalized energy vectors. The total sample size is $n = \sum_{j=1}^{m} n_j$. Now, consider a generalized ensemble whose Boltzmann probability density function is

$$\frac{1}{Z_\theta} e^{-\theta^{\mathrm{T}} u(x)}, \tag{3.1.4.49}$$

where

$$Z_\theta = \int e^{-\theta^{\mathrm{T}} u(x)} \, \mathrm{d}x \tag{3.1.4.50}$$

is the generalized configurational integral in physics or the normalization constant in statistics, and the superscript T is the transpose operator. The induced probability density function of $u(x)$ at θ is

$$\frac{1}{Z_\theta} \Omega(u) e^{-\theta^{\mathrm{T}} u}, \tag{3.1.4.51}$$

where $\Omega(u)$, formally defined as

$$\Omega(u) = \int \delta(u(x) - u) \, \mathrm{d}x, \tag{3.1.4.52}$$

is a generalized density of states, which does not depend on θ. The generalized configurational integral can also be determined from $\Omega(u)$ as

$$Z_\theta = \int \Omega(u) e^{-\theta^{\mathrm{T}} u} \, \mathrm{d}u. \tag{3.1.4.53}$$

As we have shown that the WHAM method involves constructing a histogram, $H_r(u)$, from each sample $\{u_{ri} : i = 1, \ldots, n_r\}$, which $H_r(u)$ indicates the number of observations falling into a bin about u, for example, an interval or a rectangle if $u(x)$ is one or two-dimensional. Then $\Omega(u)$ is estimated by

$$\hat{\Omega}(u) \Delta u = \frac{\sum_{r=1}^{m} H_r(u)}{\sum_{r=1}^{m} n_r \hat{Z}_{\theta_r}^{-1} e^{-\theta_r^{\mathrm{T}} u}}, \tag{3.1.4.54}$$

where the configurational integrals $(Z_{\theta_1}, \ldots, Z_{\theta_m})$ are defined by self-consistency according to Eq. 3.1.4.54

$$\hat{Z}_{\theta_k} = \sum_u \frac{\sum_{r=1}^{m} H_r(u)}{\sum_{r=1}^{m} n_r \hat{Z}_{\theta_r}^{-1} e^{(\theta_k - \theta_r)^{\mathrm{T}} u}} \quad (k = 1, \ldots, m), \tag{3.1.4.55}$$

where the summation $\sum_u$ is taken over all possible bins centered at u of size Δu. Also, the configurational integral Z_θ at any other parameter value is estimated by

$$\hat{Z}_\theta = \sum_u \frac{\sum_{r=1}^m H_r(u)}{\sum_{r=1}^m n_r \hat{Z}_{\theta_r}^{-1} e^{(\theta-\theta_r)^{\mathrm{T}} u}}. \tag{3.1.4.56}$$

We can take

$$\frac{1}{\hat{Z}_\theta} \frac{\sum_{r=1}^m H_r(u)}{\sum_{r=1}^m n_r \hat{Z}_{\theta_r}^{-1} e^{(\theta-\theta_r)^{\mathrm{T}} u}} \tag{3.1.4.57}$$

as the weight of bin u under condition θ.

Now, let $h(u)$ be a function of u, and denoted by $\langle h \rangle_\theta$ the expectation of $h(u)$. The WHAM estimate $\hat{h}_\theta$ for $\langle h \rangle_\theta$ is

$$\hat{h}_\theta = \frac{1}{\hat{Z}_\theta} \sum_u h(u) \frac{\sum_{r=1}^m H_r(u)}{\sum_{r=1}^m n_r \hat{Z}_{\theta_r}^{-1} e^{(\theta-\theta_r)^{\mathrm{T}} u}}. \tag{3.1.4.58}$$

It is interesting to note that the summation over bins in Eq. 3.1.4.58 can be equivalently expressed in terms of a weighted average over observations

$$\hat{h}_\theta = \sum_{ji} h(u_{ji}^b) F_{ji}(\theta), \tag{3.1.4.59}$$

where u_{ji}^b is a representative generalized energy of the bin containing u_{ji}, F_{ji} is the "WHAM weight" of u_{ji} that, by comparing Eqs. 3.1.4.58 and 3.1.4.59, is defined as

$$F_{ji}(\theta) = \frac{\hat{Z}_\theta^{-1}}{\sum_{r=1}^m n_r \hat{Z}_{\theta_r}^{-1} e^{(\theta-\theta_r)^{\mathrm{T}} u_{ji}^b}} = \frac{1}{\hat{Z}_\theta} e^{-\theta^T u_{ji}^b} G_{ji} \tag{3.1.4.60}$$

and

$$G_{ji} = \frac{1}{\sum_{r=1}^m n_r \hat{Z}_{\theta_r}^{-1} e^{-\theta_r^{\mathrm{T}} u_{ji}^b}} \tag{3.1.4.61}$$

is the θ-dependent component of the WHAM weight $F_{ji}(\theta)$ for each observation.

Equation 3.1.4.59 states that the expectation value of any observable can be obtained by attaching a statistical weight $F_{ji}(\theta)$ to each observation u_{ji} which depends on the bin to which it is assigned. An obvious simplification is to express the WHAM estimate of $\langle h \rangle_\theta$ and the WHAM weights (Eq. 3.1.4.60) in terms of the actual observation u_{ji} rather than their closest bin representative u_{ji}^b.

To understand binless WHAM, it is useful to introduce the concept of the measure G defined by

$$\mathrm{d}G = \Omega(u)\, \mathrm{d}u, \tag{3.1.4.62}$$

that is, $G(A) = \int_A \Omega(u)\,\mathrm{d}u$ for every measurable set A of u. Informally, this equation says that for an infinitesimal bin about u of size $\mathrm{d}u$, the weight assigned under G is $\Omega(u)\,\mathrm{d}u$. Thereafter G is called the measure of states. The probability distribution of $u(x)$, F_θ, is related to G as

$$\mathrm{d}F_\theta = \frac{1}{Z_\theta} e^{-\theta^{\mathrm{T}} u} \Omega(u)\,\mathrm{d}u = \frac{1}{Z_\theta} e^{-\theta^{\mathrm{T}} u}\,\mathrm{d}G, \tag{3.1.4.63}$$

that is

$$F_\theta(A) = \frac{1}{Z_\theta} \int_A e^{-\theta^{\mathrm{T}} u}\,\mathrm{d}G \tag{3.1.4.64}$$

for every measurable set A of u. The configurational integral can then be expressed as

$$Z_\theta = \int e^{-\theta^{\mathrm{T}} u}\,\mathrm{d}G. \tag{3.1.4.65}$$

The pooled data $\{u_{ji} : i = 1, \ldots, n_j, j = 1, \ldots, m\}$ can be regarded as an approximate sample from the mixture distribution, F_*, whose components are $(F_{\theta_1}, \ldots, F_{\theta_m})$ with proportions $(n_1/n, \ldots, n_m/n)$. F_* is related to G as

$$\mathrm{d}F_* = \left\{ \sum_{r=1}^{m} \frac{n_r}{n} Z_{\theta_r}^{-1} e^{-\theta_r^{\mathrm{T}} u} \right\} \Omega(u)\,\mathrm{d}u = \left\{ \sum_{r=1}^{m} \frac{n_r}{n} Z_{\theta_r}^{-1} e^{-\theta_r^{\mathrm{T}} u} \right\} \mathrm{d}G \tag{3.1.4.66}$$

For an infinitesimal bin about u of size $\mathrm{d}u$, the probability assigned under F_* is the expression in the curly brackets times the weight assigned under G. Dividing both sides of Eq. 3.1.4.66 by the quantity in the curly brackets gives

$$\mathrm{d}G = \left\{ \sum_{r=1}^{m} \frac{n_r}{n} Z_{\theta_r}^{-1} e^{-\theta_r^{\mathrm{T}} u} \right\}^{-1} \mathrm{d}F_*. \tag{3.1.4.67}$$

For an infinitesimal bin about u of size $\mathrm{d}u$, the weight assigned under G is the inverse of the quantity in the curly brackets times the probability assigned under F_*.

Relationship (3.1.4.67) can be used for estimating G from the pooled data by importance weighting. Recall that the pooled data form an approximate sample from F_*. Then F_* can be estimated by the empirical distribution $\hat{F}_*$ for which each observation u_{ji} is assigned the probability n^{-1}. By Eq. 3.1.4.67, the resulting estimator $\hat{G}$ is a discrete measure for which each observation u_{ji} is assigned the weight

$$\hat{G}(u_{ji}) = \frac{1}{\sum_{r=1}^{m} n_r \hat{Z}_{\theta_r}^{-1} e^{-\theta_r^{\mathrm{T}} u_{ji}}}, \tag{3.1.4.68}$$

where

$$\begin{aligned} \hat{Z}_{\theta_k} &= \sum_{j=1}^{m} \sum_{i=1}^{n_j} e^{-\theta_k^{\mathrm{T}} u_{ji}} \hat{G}(u_{ji}) \\ &= \sum_{j=1}^{m} \sum_{i=1}^{n_j} \frac{1}{\sum_{r=1}^{m} n_r \hat{Z}_{\theta_r}^{-1} e^{(\theta_k - \theta_r)^{\mathrm{T}} u_{ji}}} \quad (k = 1, \ldots, m). \end{aligned} \tag{3.1.4.69}$$

Formulas 3.1.4.68 and 3.1.4.69 provide a binless extension of Eq. 3.1.4.54 and 3.1.4.55 in WHAM.

Again, the configurational integral Z_θ at any other parameter value is estimated by

$$\begin{aligned}\hat{Z}_\theta =& \sum_{j=1}^{m}\sum_{i=1}^{n_j} e^{-\theta^{\mathrm{T}}u_{ji}}\hat{G}(u_{ji}) \\ =& \sum_{j=1}^{m}\sum_{i=1}^{n_j} \frac{1}{\sum_{r=1}^{m} n_r \hat{Z}_{\theta_r}^{-1} e^{(\theta-\theta_r)^{\mathrm{T}}u_{ji}}}. \end{aligned} \tag{3.1.4.70}$$

The expectation $\langle h\rangle_\theta$ is by definition $Z_\theta^{-1}\int h(u)e^{-\theta^{\mathrm{T}}u}\,\mathrm{d}G$ and hence estimated by

$$\begin{aligned}&\frac{1}{\hat{Z}_\theta}\sum_{j=1}^{m}\sum_{i=1}^{n_j} h(u_{ji})e^{-\theta^{\mathrm{T}}u_{ji}}\hat{G}(u_{ji}) \\ =&\frac{1}{\hat{Z}_\theta}\sum_{j=1}^{m}\sum_{i=1}^{n_j} \frac{h(u_{ji})}{\sum_{r=1}^{m} n_r \hat{Z}_{\theta_r}^{-1} e^{(\theta-\theta_r)^{\mathrm{T}}u_{ji}}}. \end{aligned} \tag{3.1.4.71}$$

Formulas 3.1.4.70 and 3.1.4.71 provide a binless extension of Eqs. 3.1.4.56 and 3.1.4.58 in WHAM.

3.1.5 Multistate Bennett Acceptance Ratio

"An alleged scientific discovery has no merit unless it can be explained to a barmaid."

– Ernest Rutherford

"So, you can never be a good scientist unless you go to bar regularly."

– Yihan Shao

The Multistate Bennett Acceptance Ratio (MBAR) method was developed by Shirts and Chodera in 2008.[51] The following derivation quite follows Ref. [52] Imaging you have carried out a series of simulations such as umbrella sampling, or replica exchange molecular dynamics simulations. Now you have K trajectories in total and each trajectory is characterized by Hamiltonian H_k and inverse temperature β_k. The trajectories unnecessarily have the same number of conformations. Instead, the number of conformations in trajectory k is N_k. Now, you mix all the samples and randomly pick one sample out of them. The probably for this sample to have coordinates $\mathbf{R}$ is

$$p_m(\mathbf{R}) = \frac{1}{N}\sum_{k=1}^{K} N_k p_k(\mathbf{R}), \tag{3.1.5.1}$$

in which $N = \sum_{k=1}^{K} N_k$ and the subscript m means mixed ensemble. $p_k(\mathbf{R})$ is the probability of finding this snapshot in trajectory k, which satisfies

$$p_k(\mathbf{R}) = {c_k}^{-1} q_k(\mathbf{R}). \tag{3.1.5.2}$$

c_k is the normalization constant. You can see that this mixed ensemble does not follow Boltzmann statistics, even if q_k does. It can be proved that if p_k is normalized, then p_m is also normalized.

The expectation of any operator $\hat{O}$ averaged over this mixed ensemble can be calculated by

$$\langle O \rangle_m = \int O(\mathbf{R}) p_m(\mathbf{R})\, \mathrm{d}\mathbf{R} \approx \frac{1}{N}\sum_{n=1}^{N} O(\mathbf{R}_n). \tag{3.1.5.3}$$

Using energy reweighting[14], we can calculate the expectation of this operator under *any* other Hamiltonian H_i and probability p_i, which can be

expressed as

$$\begin{aligned}\langle O\rangle_i &= \int O(\mathbf{R})p_i(\mathbf{R})\,\mathrm{d}\mathbf{R}\\ &= \int O(\mathbf{R})\frac{p_i(\mathbf{R})}{p_m(\mathbf{R})}p_m(\mathbf{R})\,\mathrm{d}\mathbf{R}\\ &\approx \frac{1}{N}\sum_{n=1}^{N} O(\mathbf{R}_n)\frac{p_i(\mathbf{R}_n)}{p_m(\mathbf{R}_n)}\\ &= \frac{1}{N}\sum_{n=1}^{N} O(\mathbf{R}_n){c_i}^{-1}\frac{q_i(\mathbf{R}_n)}{p_m(\mathbf{R}_n)}\\ &= \sum_{n=1}^{N} O(\mathbf{R}_n){c_i}^{-1}\frac{q_i(\mathbf{R}_n)}{\sum_{k=1}^{K} N_k p_k(\mathbf{R}_n)} \end{aligned} \tag{3.1.5.4}$$

Let $O = 1$, we find

$$1 = \sum_{n=1}^{N} {c_i}^{-1}\frac{q_i(\mathbf{R}_n)}{\sum_{k=1}^{K} N_k p_k(\mathbf{R}_n)}. \tag{3.1.5.5}$$

Since c_i does not depend on n,

$$\begin{aligned} c_i &= \sum_{n=1}^{N}\frac{q_i(\mathbf{R}_n)}{\sum_{k=1}^{K} N_k p_k(\mathbf{R}_n)}\\ &= \sum_{n=1}^{N}\frac{q_i(\mathbf{R}_n)}{\sum_{k=1}^{K} N_k {c_k}^{-1} q_k(\mathbf{R}_n)} \end{aligned} \tag{3.1.5.6}$$

In Boltzmann statistics, $q_k(\mathbf{R}) = \exp\left[-\beta_k U_k(\mathbf{R})\right]$ and $c_k = \int q_k(\mathbf{R})d\mathbf{R}$ is the partition function or the normalization constant. *Note that we have not assumed anything about the statistics of ensemble k and i. Besides, i is unnecessarily within $\{k\}$. For instance, if we run replica exchange molecular dynamics simulations at K inverse temperatures $\beta_1, \ldots, \beta_K$, β_i can be either one of these inverse temperatures or any other inverse temperature between β_1 and β_K. But extrapolation to inverse temperatures outside the range of $[\beta_K, \beta_1]$ is not recommended.*

If q_k and q_i follow Boltzmann statistics, and we define free energy $f_i = -{\beta_i}^{-1}\ln c_i$, Eq. 3.1.5.6 becomes

$$f_i = -{\beta_i}^{-1}\ln\sum_{n=1}^{N}\frac{\exp\left(-\beta_i U_i(\mathbf{R}_n)\right)}{\sum_{k=1}^{K} N_k \exp\left(\beta_k f_k - \beta_k U_k(\mathbf{R}_n)\right)}, \tag{3.1.5.7}$$

which must be solved self-consistently and can be determined up to a constant. We can fix f_1 (to 0 usually).

Again, from Eq. 3.1.5.4, we can define

$$W_{in} = {c_i}^{-1}\frac{q_i(\mathbf{R}_n)}{\sum_{k=1}^{K} N_k {c_k}^{-1} q_k(\mathbf{R}_n)}, \tag{3.1.5.8}$$

which is the weight of snapshot n in ensemble i determined by Hamiltonian H_i and temperature β_i. Specifically, for the Boltzmann statistics,

$$W_{in} = \frac{e^{-\beta_i[U_i(\mathbf{R}_n)-f_i]}}{\sum_{k=1}^{K} N_k e^{-\beta_k[U_k(\mathbf{R}_n)-f_k]}}. \tag{3.1.5.9}$$

Here, $e^{\beta_i f_i}$ in the numerator serves as a normalization factor for W_{in}, which is unknown beforehand. Practically, we can define an unnormalized weight function

$$W'_{in} = \frac{e^{-\beta_i U_i(\mathbf{R}_n)}}{\sum_{k=1}^{K} N_k e^{-\beta_k[U_k(\mathbf{R}_n)-f_k]}}. \tag{3.1.5.10}$$

As a special case, suppose we have performed a single simulation with an inverse temperature β and potential energy function $U_0(\mathbf{R})$. The expectation of an operator $\hat{X}$ under an inverse temperature β and potential energy function $U_1(\mathbf{R})$ is

$$\begin{aligned}
\langle X \rangle_1 =& \frac{\sum_n X(\mathbf{R}_n) W'_{1n}}{\sum_n W'_{1n}} \\
=& \frac{\sum_n X(\mathbf{R}_n) \frac{e^{-\beta U_1(\mathbf{R}_n)}}{N e^{-\beta[U_0(\mathbf{R}_n)-f_0]}}}{\sum_n \frac{e^{-\beta U_1(\mathbf{R}_n)}}{N e^{-\beta[U_0(\mathbf{R}_n)-f_0]}}} \\
=& \frac{\frac{1}{N}\sum_n X(\mathbf{R}_n) e^{-\beta \Delta U}}{\frac{1}{N}\sum_n e^{-\beta \Delta U}} \\
=& \frac{\left\langle X e^{-\beta \Delta U} \right\rangle_0}{\left\langle e^{-\beta \Delta U} \right\rangle_0},
\end{aligned} \tag{3.1.5.11}$$

which returns to Eq. 2.3.0.2. Similarly, Eq. 3.1.5.7 becomes

$$\exp(-\beta f_1) = \sum_{n=1}^{N} \frac{\exp(-\beta U_1(\mathbf{R}_n))}{N \exp(\beta f_0 - \beta U_0(\mathbf{R}_n))}. \tag{3.1.5.12}$$

By moving the term $\exp(-\beta f_0)$ to left hand side of this equation, we have

$$\Delta f = -\beta^{-1} \ln \sum_{n=1}^{N} \frac{1}{N} \exp(-\beta \Delta U) = -\beta^{-1} \ln \langle \exp(-\beta \Delta U) \rangle_0, \tag{3.1.5.13}$$

which is Eq. 3.1.1.9.

Ding et al[53] showed that solving Eq. 3.1.5.7 is equivalent to minimizing

$$f(b_1, b_2, \ldots, b_K) = \frac{1}{N} \sum_{k=1}^{K} \sum_{j=1}^{N_k} \ln \left\{ \sum_{l=1}^{K} \exp\left[-U_l(\mathbf{R}_j^k) - b_l \right] \right\} + \sum_{k=1}^{K} \frac{N_k}{N} b_k \tag{3.1.5.14}$$

with respect to b_k.

3.1.6 Umbrella Integration

Umbrella Integration (UI) was developed by Kästner and Thiel in 2005.[54] Usually, the Weighted Histogram Analysis Method (discussed in Sec. 3.1.4) or Multistate Bennett Acceptance Ratio (discussed in Sec. 3.1.5) are used for the postprocessing of the trajectories from umbrella sampling. There is an assumption in these methods that a global equilibrium has been reached among all the states. However, you can imagine that in umbrella sampling the explored configurational phase space is confined in a narrow region around the center of the restraining potential of each window. Each simulation sees only the landscape locally. Therefore the assumption of a global equilibrium is not valid. Amendments have been proposed. Among them is the umbrella integration method.

It can be seen from Eq. 3.1.4.40, the unbiased free energy $A_i^u(\xi)$ from the ith biased simulation is related to the biased probability $P_i^b(\xi)$ via

$$A_i^u(\xi) = -\beta^{-1} \ln P_i^b(\xi) - w_i(\xi) + F_i, \tag{3.1.6.1}$$

where $w_i(\xi) = \frac{1}{2}K_i(\xi - \xi_i)^2$ is the biasing potential of this window, and F_i is a constant to be solved for by WHAM. Instead of searching for the optimum $\{F_i\}$ for all the windows, in UI the free energy change is computed using the central idea of thermodynamic integration (discussed in Sec. 3.1.2), i.e. via

$$\Delta A_{a\to b}^u = \int_{\xi_a}^{\xi_b} \frac{\partial A^u}{\partial \xi}\, \mathrm{d}\xi. \tag{3.1.6.2}$$

Taking partial derivative on both sides of Eq. 3.1.6.1 with respect to ξ, we have

$$\frac{\partial A_i^u}{\partial \xi} = -\beta^{-1} \frac{\partial \ln P_i^b(\xi)}{\partial \xi} - \frac{\mathrm{d}w_i}{\mathrm{d}\xi}. \tag{3.1.6.3}$$

By assuming that the biased probability follows a normal distribution

$$P_i^b(\xi) = \frac{1}{\sqrt{2\pi}\sigma_i^b} \exp\left[-\frac{1}{2}\left(\frac{\xi - \overline{\xi_i^b}}{\sigma_i^b}\right)^2\right], \tag{3.1.6.4}$$

where $\overline{\xi_i^b}$ and σ_i^b of each window are calculated from the trajectory, Eq. 3.1.6.3 becomes

$$\frac{\partial A_i^u}{\partial \xi} = \beta^{-1} \frac{\xi - \overline{\xi_i^b}}{(\sigma_i^b)^2} - K_i(\xi - xi_i), \tag{3.1.6.5}$$

with a variance[55]

$$\mathrm{var}\left(\frac{\partial A_i^u}{\partial \xi}\right) = \frac{2(\xi - \overline{\xi_i^b})^2 + (\sigma_i^b)^2}{N_i \beta^2 (\sigma_i^b)^4} \tag{3.1.6.6}$$

To combine the different windows, the reaction coordinate is divided into bins that span the whole range of ξ and are independent of the windows. For

each bin centered at ξ_{bin}, the windows are combined by a weighted average as

$$\left.\frac{\partial A(\xi)}{\partial \xi}\right|_{\xi_{bin}} = \sum_{i}^{windows} p_i(\xi_{bin}) \left(\frac{\partial A_i^u(\xi)}{\partial \xi}\right)_{\xi_{bin}}, \tag{3.1.6.7}$$

with the normalized weight

$$p_i(\xi) = N_i p_i^b \xi / \sum_i N_i P_i^b(\xi). \tag{3.1.6.8}$$

N_i is the total number of steps sampled for window i. This ensures a high weight at the center of the distribution. The variance is

$$\mathrm{var}\left(\frac{\partial A^u}{\partial \xi}\right) = \sum_{i}^{windows} p_i^2 \mathrm{var}\left(\frac{\partial A_i^u}{\partial \xi}\right). \tag{3.1.6.9}$$

High-order correction to the biased distribution can be found in Ref. [56].

3.1.7 Non-Equilibrium Work

Non-Equilibrium Work for Free Energy Difference between Two States

Non-Equilibrium Work (NEW) method for equilibrium free energy calculations was proposed by Jarzynski.[57]. In 1997, Jarzynski showed

$$\langle \exp\left[-\beta W(\tau)\right]\rangle = \exp\left(-\beta \Delta A\right), \tag{3.1.7.1}$$

which is now called the Jarzynski equality. Here, W is the accumulated work along a path $\lambda(t)$ connecting the initial and final states, with $\lambda(0) = 0$ and $\lambda(\tau) = 1$, and $\Delta A = A(1) - A(0)$ the free energy difference between these two states. $\langle \cdots \rangle$ in Eq. 3.1.7.1 is an average over a series of trajectories with the initial conditions chosen according to the equilibrium Boltzmann probability in state $\lambda(0)$. The path average samples all the realizations of dynamic paths weighted by their respective path actions under the time evolution of the system with an explicitly time-dependent Hamiltonian. This equality was also obtained by Crooks for markovian and microscopically reversible dynamics.[58]

Now, we consider creating an equilibrium configuration for the state $\lambda = 0$ and then slowly changing λ from 0 to 1. As the coupling parameter advances, the system continues to sample phase space by molecular dynamics or Monte Carlo simulations, but under an explicitly time-dependent Hamiltonian. In the limit of a very slow transformation, the system will remain close to the equilibrium. The free energy difference can then be evaluated by changing λ continuously

$$\Delta A = \lim_{\tau \to \infty} \int_0^\tau \frac{\partial H\left[\mathbf{x}(t); \lambda\right]}{\partial \lambda}\bigg|_{\lambda = \lambda(t)} \dot{\lambda}(t)\, \mathrm{d}t, \tag{3.1.7.2}$$

where $\dot{\lambda}(t)$ is the time derivative of the coupling parameter λ. In Eq. 3.1.7.2, the limit of $\tau \to \infty$ ensures that the transformation is performed infinitely slowly, and thus reversibly. The right-hand side of Eq. 3.1.7.2 is the "reversible work" done to the system during the transformation.

If the system is instead transformed between the initial and final states over a finite time interval τ, the system will not be able to sample the phase space exhaustively at each value of λ, making this transformation irreversible. As the transformation proceeds, the system will be gradually driven out of equilibrium, causing hysteresis effects. From the second law of thermodynamic, it is expected that the work $W(\tau)$ performed during the nonequilibrium transformation is on average larger than or equal to the free energy difference between the two states

$$\langle W(\tau)\rangle \geq \Delta A, \tag{3.1.7.3}$$

and the difference accounts for heat-dissipation effect. The work $W(\tau)$ performed on the system is the accumulated energy cost required to change the system

$$W(\tau) = \int_0^\tau \frac{\partial H[\mathbf{x}(t);\lambda]}{\partial \lambda}\Big|_{\lambda=\lambda(t)} \dot{\lambda}(t)\,\mathrm{d}t \tag{3.1.7.4}$$

The equality in Eq. 3.1.7.3 will normally be achieved only if the transformation is infinitely slow, $\tau \to \infty$. For paths of finite length, the amount of dissipated work, $\langle W(\tau)\rangle - \Delta A \geq 0$, depends on the chosen transformation path $\lambda(t)$.

Jarzynski equality, Eq. 3.1.7.1, immediately leads to the second law in the form of Eq. 3.1.7.3 because of the Jensen's inequality, $\langle e^{-x}\rangle \geq e^{-\langle x\rangle}$. Moreover, TI and TP can be thought as the limiting cases of the nonequilibrium process. When $\tau \to \infty$ or $\dot{\lambda}(t) \to 0$, this is an infinitely slow transformation and the Eq. 3.1.7.2 is the formula of TI

$$\Delta A = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial H(\mathbf{x}, \mathbf{p}_x, \lambda)}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda \tag{3.1.7.5}$$

When $\tau \to 0$ or $\dot{\lambda}(t) \to \infty$, this is an infinitely fast transformation where the configurations will not relax and the work is simply the change in the Hamiltonian when going from the initial to the final state,

$$\lim_{\tau\to 0} W(\tau) = H(\mathbf{x}(0);\lambda=1) - H(\mathbf{x}(0);\lambda=0) \tag{3.1.7.6}$$

Substituting the Eq. 3.1.7.6 into the Eq. 3.1.7.1, the formula of TP can be recovered

$$\Delta A = -\frac{1}{\beta} \ln \langle \exp[-\beta \Delta H(\mathbf{x}, \mathbf{p}_x)]\rangle_0\,, \tag{3.1.7.7}$$

In Ref. [58], Crooks showed that the distributions of work values from the forward and the backward paths satisfy a relation that is central to the histogram methods in free energy calculations

$$\frac{p_f[w = W(\tau)]}{p_b[w = -\underline{W}(\tau)]} = \exp[\beta(w - \Delta A)], \tag{3.1.7.8}$$

where $p_f[w = W(\tau)]$ and $p_b[w = -\underline{W}(\tau)]$ are the probability densities of the work values in the forward and the reverse transformations (with a sign change for the work in the reverse path). Both are normalized, i.e., $\int p_f(w)\,\mathrm{d}w = \int p_b(w)\,\mathrm{d}w = 1$. It is noted that Jarzynski equality Eq. 3.1.7.1 follows from Eq. 3.1.7.8 simply by integration over w because the probability densities are normalized to 1:

$$\int p_f(W) e^{-\beta W}\,\mathrm{d}W = \int p_b(W) e^{-\beta \Delta A}\,\mathrm{d}W, \tag{3.1.7.9}$$

Because of the normalization condition, the right-hand side is equal to $\exp(-\beta\Delta A)$, and Jarzynski equality follows.

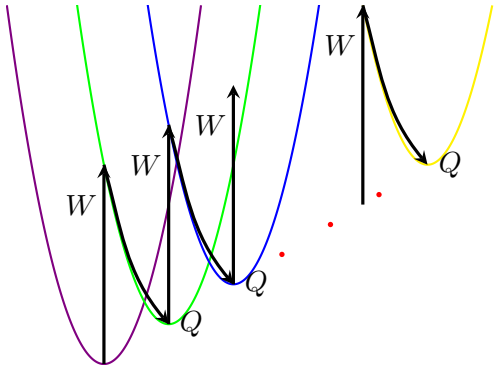

Figure 3.3: The accumulation of work and heat along a nonequilibrium trajectory. The work is defined as the energy change when the coupling parameter switches from λ_i to λ_{i+1} with the coordinates fixed, while the dissipated heat is defined as the energy relaxation when the coordinate change with the coupling parameter fixed.

Following the Crooks Fluctuation Theorem (CFT),[58] Bennett acceptance ratio can be applicable to nonequilibrium calculations. This approach was combined with a maximum likelihood estimate, and accurate free energy differences were obtained.[44] In this approach, ΔA is calculated via

$$\sum_{i=1}^{n_F} \frac{1}{1+\exp\left[\beta(M+W_i-\Delta A)\right]} = \sum_{j=1}^{n_R} \frac{1}{1+\exp\left[-\beta(M+W_j-\Delta A)\right]}, \tag{3.1.7.10}$$

where n_F and n_R are the numbers of the forward and reverse transformations respectively, W_i and W_j are the work of forward and reverse measurements respectively, and $M = \beta^{-1}\ln(n_F/n_R)$. The corresponding statistical variance of ΔA, σ^2, is calculated using Eq. 10 in Ref. [44].

Non-Equilibrium Work for Free Energy Profiles

When calculating the free energy profile in a pulling experiment, the Jarzynski equality is no longer straightforwardly applicable, because it relates the nonequilibrium work to free energy differences at different times, not positions along a predefined reaction coordinate. In order to surmount this difficulty, Hummer and Szabo extended the Jarzynski equality by measuring force/extension along pulling.[59]

Let us begin with a system of which the phase-space density evolves according to a Liouville-type equation:

$$\frac{\partial f(\mathbf{x},t)}{\partial t} = \mathscr{L}_t f(\mathbf{x},t). \tag{3.1.7.11}$$

$\mathscr{L}_t$ is an explicitly time-dependent evolution operator that has the Boltzmann distribution as a stationary solution, $\mathscr{L}_t e^{-\beta\mathscr{H}(\mathbf{x},t)} = 0$, where $\mathscr{H}(\mathbf{x},t)$ is a time-dependent Hamiltonian. For diffusive dynamics on a potential $V(\mathbf{x},t)$, the time evolution is governed by $\mathscr{L}_t = D\nabla e^{-\beta V(\mathbf{x},t)}\nabla e^{\beta V(\mathbf{x},t)}$, where D is the diffusion coefficient and $\nabla = \partial/\partial\mathbf{x}$. Now consider the unnormalized Boltzmann distribution at time t,

$$p(\mathbf{x},t) = \frac{e^{-\beta\mathscr{H}(\mathbf{x},t)}}{\int e^{-\beta\mathscr{H}(\mathbf{x}',0)}\,\mathrm{d}\mathbf{x}'}. \tag{3.1.7.12}$$

Because this distribution is stationary ($\mathscr{L}_t p = 0$), and because $\partial p/\partial t = -\beta(\partial\mathscr{H}/\partial t)p$, it follows that the above $p(\mathbf{x},t)$ is a solution of the sink equation

$$\frac{\partial p}{\partial t} = \mathscr{L}_t p - \beta\frac{\partial\mathscr{H}}{\partial t}p, \tag{3.1.7.13}$$

of which the solution, starting from an equilibrium distribution at time $t = 0$, can be expressed as a path integral by using the Feynman-Kac theorem. Equating these two differential solutions immediately gives:

$$\frac{e^{-\beta\mathscr{H}(\mathbf{x},t)}}{\int e^{-\beta\mathscr{H}(\mathbf{x}',0)}d\mathbf{x}'} = \left\langle \delta(\mathbf{x}-\mathbf{x}_t)\exp\left[-\beta\int_0^t \frac{\partial\mathscr{H}}{\partial t'}(\mathbf{x}_{t'},t')\,\mathrm{d}t'\right]\right\rangle. \tag{3.1.7.14}$$

The average $\langle\cdots\rangle$ is over an ensemble of trajectories starting from the equilibrium distribution at $t = 0$ and evolving according to Eq. 3.1.7.11. Each trajectory is weighted with the Boltzmann factor of the external work w_t done to the system,

$$w_t = \int_0^t \frac{\partial\mathscr{H}}{\partial t'}(\mathbf{x}_{t'},t')\,\mathrm{d}t'. \tag{3.1.7.15}$$

Integrating on both sides of Eq. 3.1.7.14 with respect to $\mathbf{x}$, we obtain Jarzynski equality

$$e^{-\beta\Delta G(t)} \equiv \frac{\int e^{-\beta\mathscr{H}(\mathbf{x},t)}\,\mathrm{d}\mathbf{x}}{\int e^{-\beta\mathscr{H}(\mathbf{x},0)}d\mathbf{x}} = \left\langle e^{-\beta w_t}\right\rangle \tag{3.1.7.16}$$

between the Boltzmann-averaged work w_t and the equilibrium free energy difference $\Delta G(t)$ between times t and 0.

In a single-molecule pulling experiment, e.g. using atomic force microscope (AFM), the sample is moved at a constant speed v relative to the cantilever with spring constant k. The position $z_t = vt + \delta z_t$ of the cantilever tip with respect to the sample is recorded, where δz_t is the displacement of the cantilever tip. It can be described by a Hamiltonian $\mathscr{H}(\mathbf{x},t) = \mathscr{H}_0(\mathbf{x}) + k(z-vt)^2/2$, where $\mathscr{H}_0(\mathbf{x})$ is the Hamiltonian of the resting, unperturbed system, and $z = z(\mathbf{x})$. Substituting this Hamiltonian into Eq. 3.1.7.14, multiplying both sides by $\delta[z - z(\mathbf{x})]$, integrating over all

$\mathbf{x}$, and finally taking the logarithm, we have:

$$\begin{aligned} G_0(z) \equiv & -\beta^{-1} \ln \frac{\int \delta[z - z(\mathbf{x})] e^{-\beta \mathscr{H}_0(\mathbf{x})} \, \mathrm{d}\mathbf{x}}{\int e^{-\beta \mathscr{H}(\mathbf{x},0)} \, \mathrm{d}\mathbf{x}} \\ = & -\beta^{-1} \ln \left\langle \delta(z - z_t) e^{-\beta \Delta w_t} \right\rangle, \end{aligned} \tag{3.1.7.17}$$

where $G_0(z)$ is the unperturbed free energy profile along the pulling coordinate z, and Δw_t is the external work minus the instantaneous biasing potential, $\Delta w_t = w_t - k(z_t - vt)^2/2 = kv(vt^2/2 - \int_0^t z_{t'} \, \mathrm{d}t') - k(z_t - vt)^2/2$. The expression can be obtained by integrating the power over time. At any moment t', the force acting on the system is $f = k(vt' - z_{t'})$. Correspondingly, the power $p = f \cdot v$. Therefore, $w_t = \int p \, \mathrm{d}t' = \int_0^t kv(vt' - z_{t'}) \, \mathrm{d}t' = kv(vt^2/2 - \int_0^t z_{t'} \, \mathrm{d}t')$. At time $t = 0$, the trajectories are started from points z_0 drawn from a Boltzmann distribution corresponding to Hamiltonian $\mathscr{H}(\mathbf{x}, 0) = \mathscr{H}_0(\mathbf{x}) + kz^2/2$, which is *NOT* $\mathscr{H}_0(\mathbf{x})$.

At each time slice t, one can in principle obtain an estimate of the whole free energy surface. In practice with finite number of trajectories, at any give time t, only a narrow region around the equilibrium position $z = vt$ will be sampled adequately. Thus, an average over several time slices and repeated trajectories is required to obtain an optimal estimate of the free energy profile. At every time slice t, one obtains an ensemble of positions z_t and corresponding w_ts. The position z_t are binned, and the corresponding histogram values are incremented by $e^{-\beta w_t}$. The complete free energy profile $G_0(z)$ can be reconstructed by adapting the weighted histogram method:

$$G_0(z) = -\beta^{-1} \ln \frac{\sum_t \frac{\langle \delta(z - z_t) \exp(-\beta w_t) \rangle}{\langle \exp(-\beta w_t) \rangle}}{\sum_t \frac{\exp[-\beta u(z,t)]}{\langle \exp(-\beta w_t) \rangle}}, \tag{3.1.7.18}$$

where the sum is over time slices t and $u(z,t) = k(z - vt)^2/2$ is the time dependent biasing potential. As in the weighted histogram method, this procedure can be refined by making Eq. 3.1.7.18 self-consistent through replacement of $\langle \exp(-\beta w_t) \rangle$ with

$$\exp[-\beta \Delta G(t)] = \frac{\int \exp\{-\beta [u(z,t) + G_0(z)]\} \, \mathrm{d}z}{\int \exp\{-\beta [u(z,0) + G_0(z)]\} \, \mathrm{d}z}, \tag{3.1.7.19}$$

thus requiring an iterative solution for $\Delta G(t)$. Note that Eq. 3.1.7.18 can be rewritten as

$$G_0(z) = -\beta^{-1} \ln \frac{\sum_t \frac{\langle \delta(z - z_t) \exp(-\beta \Delta w_t) \rangle \exp[-\beta u(z,t)]}{\langle \exp(-\beta w_t) \rangle}}{\sum_t \frac{\exp[-\beta u(z,t)]}{\langle \exp(-\beta w_t) \rangle}}, \tag{3.1.7.20}$$

which is the natural logarithm of the weighted average of $\langle \delta(z - z_t) \exp(-\beta \Delta w_t) \rangle$ over all the time slices with $\frac{\exp[-\beta u(z,t)]}{\langle \exp(-\beta w_t) \rangle}$ being the weight.

3.1.8 Transition-Based Reweighting Analysis Method

Transition-Based Reweighting Analysis Method (TRAM), which relies on the maximum likelihood analysis of the thermodynamic and kinetic information, was developed by Frank Noé and coworkers in 2014.[60]. It incorporates WHAM with Markov state model, and avoids the weakness in both methods. In WHAM, a global equilibrium among all the thermodynamic states must be reached. For the Markov state model (MSM), the kinetic information can be extracted from only one thermodynamic state. In contrast, TRAM is a class of estimators that (1) take the statistical weights of samples at different thermodynamic states into account, and (2) exploit transitions observed in the sampled trajectories, without assuming that these trajectories are sampled from equilibrium.

Let us assume that there are K molecular dynamics (MD) or Markov chain Monte Carlo (MCMC) simulations have been performed, each in a specific thermodynamic state (Hamiltonian, temperature, etc) indexed by $k \in \{1, \ldots, K\}$. For simulations with varying thermodynamic state in a single trajectory, with replica-exchange simulation being a typical example, each contiguous sequence is treated as a separated trajectory at one of the K thermodynamic states. We further assume the configuration space (that has been visited by the simulations) is discretized into cells indexed by $i, j \in \{1, \ldots, n\}$.

Similar to the WHAM analysis, the unbiased probability, π_i, and the biased probability under thermodynamic state k, $\pi_i^{(k)}$ are related by a known and constant bias factor $\gamma_i^{(k)}$

$$\pi_i^{(k)} = f^{(k)} \pi_i \gamma_i^{(k)}, \tag{3.1.8.1}$$

$$f^{(k)} = \frac{1}{\sum_l \pi_l \gamma_l^{(k)}}, \tag{3.1.8.2}$$

where $f^{(k)}$ is a normalization constant. Thus, the bias is multiplicative in probability or additive in the potential. As we have shown in section 3.1.4, the WHAM estimator can be derived by maximizing the likelihood

$$L_{\mathrm{WHAM}} = \prod_k \prod_i (\pi_i^{(k)})^{N_i^{(k)}} \tag{3.1.8.3}$$

with an implied assumption that every count $N_i^{(k)}$ is independently drawn from the biased distribution $\pi_i^{(k)}$.

The maximum likelihood Markov model is the transition matrix $\mathbf{P} = (p_{ij})$ between n discrete configuration states, that maximizes the likelihood of the observed transitions between these states. The likelihood of a Markov model is a product of all transition probabilities corresponding to the observed trajectory. To obtain a reversible Markov state model, this likelihood

is maximized under the constraints of detailed balance with respect to the equilibrium distribution $\boldsymbol{\pi}$

$$L_{\mathrm{MSM}} = \prod_i \prod_j p_{ij}^{c_{ij}}, \tag{3.1.8.4}$$

$$s.t. \quad \pi_i p_{ij} = \pi_j p_{ji} \quad \text{for all } i, j, \tag{3.1.8.5}$$

where c_{ij} is the number of times the trajectories were observed in state i at time t and in state j at a later time $t + \tau$, where τ is the lag time at which the Markov model is estimated.

In TRAM, WHAM and MSM are combined as follows: every trajectory at thermodynamic state k is treated as a Markov chain with the configuration-state transition counts $c_{ij}^{(k)}$, without assuming that every count is sampled from global equilibrium. In contrast to Markov models, we exploit the fact that equilibrium probabilities can be reweighted between different thermodynamic states via Eqs. 3.1.8.1 and 3.1.8.2. The resulting likelihood of all $\mathbf{P}^{(k)}$ and $\boldsymbol{\pi}$, based on simulations at all thermodynamic states can be formulated as

$$L_{\mathrm{TRAM}} = \prod_k \prod_i \prod_j (p_{ij}^{(k)})^{c_{ij}^{(k)}}, \tag{3.1.8.6}$$

$$s.t. \quad \pi_i^{(k)} p_{ij}^{(k)} = \pi_j^{(k)} p_{ji}^{(k)} \quad \text{for all } i, j, k. \tag{3.1.8.7}$$

Here, $\mathbf{P}^{(k)} = (p_{ij}^{(k)})$ is the Markov transition matrix at thermodynamic state k, and $c_{ij}^{(k)}$ are the number of transitions observed at that simulation condition. $\boldsymbol{\pi}^{(k)}$ is the vector of equilibrium probabilities of discrete states at each thermodynamic state. *Because each Markov model* $\mathbf{P}^{(k)}$ *must have the distribution* $\boldsymbol{\pi}^{(k)}$ *as a stationary distribution, all Markov models are coupled too, which makes the maximization of the TRAM likelihood Eqs. 3.1.8.6 and 3.1.8.7 difficult, and it can neither be achieved by WHAM, nor by existing MSM estimators.*

Taking natural logarithm on the TRAM likelihood, we find

$$\ln L_{\mathrm{TRAM}} = \sum_{k=1}^{K} \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}^{(k)} \ln p_{ij}^{(k)}, \tag{3.1.8.8}$$

with constraints

$$\pi_i \gamma_i^{(k)} p_{ij}^{(k)} = \pi_j \gamma_j^{(k)} p_{ji}^{(k)} \quad \text{for all } i, j, k, \tag{3.1.8.9}$$

which is from Eqs. 3.1.8.1 and 3.1.8.7 with $f^{(k)}$ canceled. In addition, $\mathbf{P}^{(k)}$ and $\boldsymbol{\pi}$ should satisfy the normalization conditions

$$\sum_j p_{ij}^{(k)} = 1 \quad \forall i, k, \tag{3.1.8.10}$$

$$\sum_j \pi_j = 1. \tag{3.1.8.11}$$

The normalization of $\boldsymbol{\pi}^{(k)}$ is naturally satisfied due to Eqs. 3.1.8.1 and 3.1.8.2. Due to the existence of constraints, the numbers of free variables are $n-1$ for $\boldsymbol{\pi}$ and $n(n-1)/2$ for $\mathbf{P}^{(k)}$.

Using the Lagrange duality theory, it can be shown that the optimal solution of the discrete TRAM problem above fulfills the following two conditions

$$\sum_k \sum_j \frac{(c_{ij}^{(k)} + c_{ji}^{(k)})\gamma_i^{(k)}\pi_i\nu_j^{(k)}}{\gamma_i^{(k)}\pi_i\nu_j^{(k)} + \gamma_j^{(k)}\pi_j\nu_i^{(k)}} = \sum_k \sum_j c_{ji}^{(k)}, \tag{3.1.8.12}$$

$$\sum_j \frac{(c_{ij}^{(k)} + c_{ji}^{(k)})\gamma_j^{(k)}\pi_j}{\gamma_i^{(k)}\pi_i\nu_j^{(k)} + \gamma_j^{(k)}\pi_j\nu_i^{(k)}} = 1 \tag{3.1.8.13}$$

where $\nu_i^{(k)}$ are unknown Lagrange multipliers. To numerically solve the discrete TRAM problem, an initial guess for $\boldsymbol{\pi}$ and $\mathbf{v}^{(k)}$ can be made as

$$\pi_i^{init} := 1/n, v_i^{(k),init} := \sum_j c_{ij}^{(k)}, \tag{3.1.8.14}$$

and the following equations must be solved iteratively until $\boldsymbol{\pi}$ is converged:

$$v_i^{(k),new} := v_i^{(k)} \sum_j \frac{(c_{ij}^{(k)} + c_{ji}^{(k)})\gamma_i^{(k)}\pi_j}{\gamma_i^{(k)}\pi_i\nu_j^{(k)} + \gamma_j^{(k)}\pi_j\nu_i^{(k)}}, \tag{3.1.8.15}$$

$$\pi_i^{new} := \frac{\sum_{k,j} c_{ji}^{(k)}}{\sum_{k,j} \frac{(c_{ij}^{(k)}+c_{ji}^{(k)})\gamma_i^{(k)}\nu_j^{(k)}}{\gamma_i^{(k)}\pi_i\nu_j^{(k)}+\gamma_j^{(k)}\pi_j\nu_i^{(k)}}}. \tag{3.1.8.16}$$

3.2 Approximate Methods

3.2.1 Molecular Mechanics/Poisson-Boltzmann Surface Area

The following derivation follows Ref. [61]. The Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) method is often used in the calculations of binding free energy of a substrate to a receptor. The standard binding free energy for a reaction between a receptor (A) and a substrate (B)

$$\mathrm{A} \quad + \quad \mathrm{B} \rightleftharpoons \mathrm{AB}$$

is expressed as a ratio of configuration integrals

$$\begin{aligned}\Delta G^0_{AB} =& -\beta^{-1} \ln\left(\frac{C^0}{8\pi^2}\cdot\frac{Z_{N,AB}Z_{N,0}}{Z_{N,A}Z_{N,B}}\right) + P^0\left\langle\Delta V_{AB}\right\rangle \\ =& -\beta^{-1}\ln\left(\frac{C^0}{8\pi^2}\cdot\frac{\frac{Z_{N,AB}}{Z_{N,0}}}{\frac{Z_{N,A}}{Z_{N,0}}\frac{Z_{N,B}}{Z_{N,0}}}\right) + P^0\left\langle\Delta V_{AB}\right\rangle, \end{aligned} \tag{3.2.1.1}$$

where R is the gas constant, T is the temperature, C^0 is the standard state concentration (1 M), N is the number of solvent molecules, and $P^0\left\langle\Delta V_{AB}\right\rangle$ is the pressure-volume work associated with changing the system size after the association of two species into one complex. For water solution at 1 atm, the last term is negligibly small. There are no mass-dependent terms in Eq. 3.2.1.1, which is a direct result of equal kinetic contribution to the partition function of the bound and the free species. The configuration integral of the receptor, A, in solution is

$$Z_{N,A} = \int e^{-\beta U(r_A,r_S)}\,\mathrm{d}r_A\,\mathrm{d}r_S, \tag{3.2.1.2}$$

where $U(r_A, r_S)$ is the potential energy as a function of all solute coordinates, r_A, and solvent coordinates, r_S, and β is the reciprocal of the product of the Boltzmann constant and temperature. The total potential energy can be decomposed into $U(r_A)+U(r_S)+\Delta U(r_A, r_S)$. Similar for B, the substrate. For pure solvent, the configuration integral is

$$Z_{N,0} = \int e^{-\beta U(r_S)}\,\mathrm{d}r_S. \tag{3.2.1.3}$$

The ratio of configuration integrals in Eq. 3.2.1.1 can be simplified with an implicit solvent approximation, as

$$\begin{aligned}\frac{Z_{N,A}}{Z_{N,0}} = Z_A =& \frac{\int e^{-\beta U(r_A)}\left\{\int e^{-\beta\Delta U(r_A,r_S)}e^{-\beta U(r_S)}\,\mathrm{d}r_S\right\}\mathrm{d}r_A}{\int e^{-\beta U(r_S)}\,\mathrm{d}r_S} \\ =& \int e^{-\beta[U(r_A)+W(r_A)]}\,\mathrm{d}r_A, \end{aligned} \tag{3.2.1.4}$$

where

$$W(r_A) = -\beta^{-1} \ln \left(\frac{\int e^{-\beta \Delta U(r_A, r_S)} e^{-\beta U(r_S)} \, \mathrm{d}r_S}{\int e^{-\beta U(r_S)} \, \mathrm{d}r_S} \right) \tag{3.2.1.5}$$

is the solvation free energy of the receptor A at fixed coordinate r_A. Analogous equations hold for the complex and substrate.

For the complex, we define the position (translational degrees of freedom) and orientation (rotational degrees of freedom) of the substrate with respective to the receptor as $\delta_B \equiv (x_1, x_2, x_3, \xi_1, \xi_2, \xi_3)$. Generally, these degree-of-freedom is very limited. The complex configuration integral is

$$Z_{AB} = \int e^{-\beta [U(r_A, r_{B'}, \delta_B) + W(r_A, r_{B'}, \delta_B)]} \, \mathrm{d}r_A \, \mathrm{d}r_{B'} \, \mathrm{d}\delta_B, \tag{3.2.1.6}$$

where $r_{B'}$ represents the remaining internal degrees of freedom of the bound substrate and δ_B spans conformations where A and B form a complex. Then we assume that the translational and rotational motions of the substrate in the bound state are not strongly coupled with the other degrees of freedom, and we decompose the potential and solvation energies as (*so weird!*)

$$\begin{aligned} &U\left(r_A, r_{B'}, \delta_B\right) + W\left(r_A, r_{B'}, \delta_B\right) \\ &\quad \approx U_1(\delta_B) + W_1(\delta_B) + U_2(r_A, r_{B'}) + W_2(r_A, r_{B'}). \end{aligned} \tag{3.2.1.7}$$

We further assume that the residual translational and rotational motions of the substrate are uncorrelated. Therefore

$$U_1(\delta_B) \approx U(x_1, x_2, x_3) + U(\xi_1, \xi_2, \xi_3), \tag{3.2.1.8}$$

and

$$W_1(\delta_B) \approx W(x_1, x_2, x_3) + W(\xi_1, \xi_2, \xi_3). \tag{3.2.1.9}$$

Now, Eq. 3.2.1.1 can bre written as

$$\Delta G^0_{AB} = -\beta^{-1} \ln \left[\frac{C^0 Z^{trans}_{B'} Z^{rot}_{B'} Z_{AB'}}{8\pi^2 Z_A Z_B} \right], \tag{3.2.1.10}$$

where

$$Z^{trans}_{B'} = \int e^{-\beta [U(x_1, x_2, x_3) + W(x_1, x_2, x_3)]} \, \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}x_3 \tag{3.2.1.11}$$

and

$$Z^{rot}_{B'} = \int e^{-\beta [U(\xi_1, \xi_2, \xi_3) + W(\xi_1, \xi_2, \xi_3)]} \, \mathrm{d}\xi_1 \, \mathrm{d}\xi_2 \, \mathrm{d}\xi_3. \tag{3.2.1.12}$$

As a first-order approximation, we assume that the energetic landscape of each species has an energy and a volume (entropy),

$$Z_A = \int e^{-\beta [U(r_A) + W(r_A)]} \, \mathrm{d}r_A \approx Z^{int}_A e^{-\beta \langle E_A \rangle}, \tag{3.2.1.13}$$

where $\langle E_A \rangle = \langle U(r_A) + W(r_A) \rangle$. We further assume (*how many approximations we have taken!*) that $Z_A^{int} Z_B^{int} \approx Z_{AB}^{int}$, then

$$\Delta G_{AB}^0 = -\beta^{-1} \ln \left(\frac{C^0 Z_{B'}^{trans} Z_{B'}^{rot}}{8\pi^2} \right) + \left(\langle E_{AB'} \rangle - \langle E_A \rangle - \langle E_B \rangle \right). \quad (3.2.1.14)$$

The bound substrate's translational configuration integral, $Z_{B'}^{trans}$, can be conceptually linked to the volume of space that its center of mass occupies through the simulation. The effective volume was measured based on the assumption that the translational motion is restrained by three harmonic potential. By solving eigenstates of the center-of-mass covariance matrix, the eigenvalues describe the variance Δx_i^2 along each principal axis. Thus, the translational configuration integral can be calculated as

$$\begin{aligned} Z_{B'}^{trans} =& \int e^{\left(-k_1 \Delta x_1^2 / 2k_B T\right)} \mathrm{d}x_1 \int e^{\left(-k_2 \Delta x_2^2 / 2k_B T\right)} \mathrm{d}x_2 \int e^{\left(-k_3 \Delta x_3^2 / 2k_B T\right)} \mathrm{d}x_3 \\ =& (2\pi)^{3/2} \left(\left\langle \Delta x_1^2 \right\rangle \left\langle \Delta x_2^2 \right\rangle \left\langle \Delta x_3^2 \right\rangle \right)^{1/2} \end{aligned} \quad (3.2.1.15)$$

where

$$k_i = \frac{k_B T}{\left\langle \Delta x_i^2 \right\rangle}. \quad (3.2.1.16)$$

The rotational configuration integral can be accounted in a similar manner.

4

Evaluation of Reliability

"A theory is something nobody believes, except the person who made it. An experiment is something everybody believes, except the person who made it."

– Albert Einstein,

4.1 Overlap Matrix

Overlap matrix proposed by Mobley et. al.,[62] can be used to essentially measures the magnitude of the phase space overlap. For example, after MBAR method is used to analyze the US simulations or a series of alchemical window simulations, the overlap matrix can be used to examine the reliabilities of MBAR calculations. The formula about the overlap matrix is shown as follows:

For the US simulations, with the weight of the lth configuration in the ith biased simulation appearing in the tth simulation defined as

$$w_t(\mathbf{x}_{i,l}) = \frac{e^{-\beta\left[W_t(\mathbf{x}_{i,l}) - f_t^{(b)}\right]}}{\sum_{k=1}^{S} N_k e^{-\beta\left[W_k(\mathbf{x}_{i,l}) - f_k^{(b)}\right]}}, \tag{4.1.0.1}$$

the elements of the $S \times S$ overlap matrix are[62]

$$\begin{aligned} O_{tt'} &= \sum_{i=1}^{S}\sum_{l=1}^{N_i} N_t w_t(\mathbf{x}_{i,l}) w_{t'}(\mathbf{x}_{i,l}) \\ &= \sum_{i=1}^{S}\sum_{l=1}^{N_i} \frac{N_t e^{-\beta\left[W_t(\mathbf{x}_{i,l}) - f_t^{(b)}\right]} e^{-\beta\left[W_{t'}(\mathbf{x}_{i,l}) - f_{t'}^{(b)}\right]}}{\left\{\sum_{k=1}^{S} N_k e^{-\beta\left[W_k(\mathbf{x}_{i,l}) - f_k^{(b)}\right]}\right\}^2}. \end{aligned} \tag{4.1.0.2}$$

Consecutive windows should have substantial overlap with the diagonal and the first off-diagonal elements no smaller than 0.03 as recommended[62].

For a series of alchemical window simulations, it is a $K \times K$ matrix with entries

$$O_{ij} = \sum_{n=1}^{N} \frac{N_i p_i(x_n)}{\sum_{k=1}^{K} N_k p_k(x_n)} \frac{p_j(x_n)}{\sum_{k=1}^{K} N_k p_k(x_n)}, \tag{4.1.0.3}$$

where $p_i(x_n) = e^{\beta G_i - \beta U_i(x_n)}$ is the probability of sample x_n occurring when simulation state i and N samples are collected with N_1 samples from $p_1(x)$ distribution, N_2 samples from $p_2(x)$ distribution, and so on. K is the total number of the states. O_{ij} can be interpreted as the average probability of a sample generated in state j being observed in the ith state. The average is computed over samples collected from all the K states, not just the samples from state j. Therefore O_{ij} is a measure of the overlap in the phase space of state i and j. The larger the better. The largest eigenvalue is 1. Similarly, consecutive windows should have substantial overlap with the diagonal and the first off-diagonal elements no smaller than 0.03 as recommended[62].

4.2 Π Metric for Neglected-tail Bias Model

The neglected-tail bias model is developed by Kofke and coworkers for the estimate of bias in free energy via thermodynamic perturbation or the Jarzynski equality[63, 64]. Here, we will take the latter as an example. Let $p_A(W)$ and $p_B(W)$ be the distributions/probabilities of work samples W in a forward ($A \to B$) and a backward ($B \to A$) nonequilibrium conversions. The Jarzynski equality discussed in 3.1.7 shows that the free energy difference is given by

$$\exp\left(-\beta \Delta A\right) = \int_{-\infty}^{\infty} p_A(W) e^{-\beta W} \, \mathrm{d}W, \tag{4.2.0.1}$$

or

$$\exp\left(+\beta \Delta A\right) = \int_{-\infty}^{\infty} p_B(W) e^{+\beta W} \, \mathrm{d}W. \tag{4.2.0.2}$$

These two distributions are related

$$p_A(W) e^{-\beta W} = p_B(W) e^{-\beta \Delta A}. \tag{4.2.0.3}$$

The neglected-tail bias model asserts that *all* of the bias is due the neglect of contributions below a particular value of W, designated as W^*, such that no sampling is contributed below this value, and *perfect sampling* is achieved for $W > W^*$.

It can be easily imagined that W^* depends on the sample size. When more sampling is performed, more likely a more negative value of W^* will

be encountered. Given $p_A(W)$, the probability distribution for W^* being observed once within M samples is

$$P_A^*(W^*) = Mp_A(W^*)[1 - C_A(W^*)]^{M-1}, \tag{4.2.0.4}$$

where $C_A(W)$ is the cumulative distribution function defined as

$$C_A(W^*) = \int_{-\infty}^{W^*} p_A(W)\,\mathrm{d}W \tag{4.2.0.5}$$

and

$$1 - C_A(W^*) = \int_{W^*}^{\infty} p_A(W)\,\mathrm{d}W. \tag{4.2.0.6}$$

The bias can be written as

$$\begin{aligned} B(M) =& \langle \Delta A(M) \rangle - \Delta A \\ =& -\beta^{-1} \ln \left[\frac{1}{M} \left(e^{-\beta W^*} + (M-1) \times \frac{\int_{W^*}^{\infty} \mathrm{d}W e^{-\beta W} p_A(W)}{1 - C_A(W^*)} \right) \right] - \Delta A \\ =& -\beta^{-1} \ln \left[\frac{1}{M} \left(e^{-\beta W_{dis}^*} + (M-1) \times \frac{1 - C_B(W^*)}{1 - C_A(W^*)} \right) \right], \end{aligned} \tag{4.2.0.7}$$

where

$$W_{dis} = W - \Delta A \tag{4.2.0.8}$$

is the dissipated work. For the third equality, we have employed Eq. 4.2.0.3. The bias can be evaluated as the average of the inaccuracy over the distribution of W^*

$$B(M) = -\beta^{-1} \int_{-\infty}^{\infty} \mathrm{d}W^* P_A^*(W^*) \ln \left[\frac{1}{M} \left(e^{-\beta W_{dis}^*} + (M-1) \times \frac{1 - C_B(W^*)}{1 - C_A(W^*)} \right) \right]. \tag{4.2.0.9}$$

Alternatively, it can be estimated for a single value of W^*, for instance the mode $\hat{W}^*$ of P_A^*. It follows from

$$\left. \frac{\mathrm{d} \ln P_A(W^*)}{\mathrm{d}W^*} \right|_{W = \hat{W}^*} = 0 \tag{4.2.0.10}$$

and is given by the solution of

$$\left. \frac{\mathrm{d} \ln p_A(W)}{\mathrm{d}W} \right|_{W = \hat{W}^*} = \frac{(M-1) p_A(\hat{W}^*)}{1 - C_A(\hat{W}^*)}. \tag{4.2.0.11}$$

or, with a small approximation

$$\left. \frac{\mathrm{d} \ln p_A(W)}{\mathrm{d}W} \right|_{W = \hat{W}^*} = (M-1) p_A(\hat{W}^*). \tag{4.2.0.12}$$

Assuming W^* has a Gaussian distribution centered at $\hat{W}^*$ as

$$P_A^*(W^*) \approx \frac{1}{\sqrt{2\pi}\sigma^*} \exp\left[-(W^* - \hat{W}^*)^2/2\sigma^{*2}\right] \tag{4.2.0.13}$$

in which, using Eq. 4.2.0.4 and the property in Eq. 4.2.0.11,

$$\frac{1}{\sigma^{*2}} = -\frac{\mathrm{d}^2 \ln P_A^*}{\mathrm{d}\, W^{*2}} = -\left[\frac{\partial^2 \ln p_A}{\partial W^2} - \frac{M}{M-1}\left(\frac{\partial \ln p_A}{\partial W}\right)^2\right]\Bigg|_{W=\hat{W}^*}. \tag{4.2.0.14}$$

Corresponding relations can be given for the bias of the $B \to A$ work process. Thus from the work distributions the expected performance of a simulation of given sampling length can be predicted.

Next, let us look at the work distributions. We assume a Gaussian-work distribution

$$p_A(W) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-(W - \bar{W})^2/2\sigma^2\right]. \tag{4.2.0.15}$$

From Eq. 4.2.0.3, it can be found that $p_B(W)$ is also a Gaussian with the same variance but shifted by $\beta\sigma^2$

$$p_B(W) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-(W - \bar{W} + \beta\sigma^2)^2/2\sigma^2\right]. \tag{4.2.0.16}$$

The free energy difference is

$$\Delta A = \bar{W} - \beta\sigma^2. \tag{4.2.0.17}$$

The cumulative distribution functions are

$$C_A(W) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\bar{W} - W}{\sqrt{2}\sigma}\right), \tag{4.2.0.18}$$

$$C_B(W) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\bar{W} - \beta\sigma^2 - W}{\sqrt{2}\sigma}\right), \tag{4.2.0.19}$$

where $\mathrm{erfc}(x)$ is the complementary error function.

By taking Eq. 4.2.0.15 into Eq. 4.2.0.12, the mode of the least-work distribution is approximately

$$\hat{W}^* = \bar{W} - \sigma\sqrt{\mathbf{W}_L\left[\frac{1}{2\pi}(M-1)^2\right]}, \tag{4.2.0.20}$$

where $\mathbf{W}_L(x)$ is the Lambert W function, defined as the solution to $x = we^w$. With a slight approximation to Eq. 4.2.0.7 as

$$B(M) = -\beta^{-1} \ln\left[1 - C_B(W^*)\right] \tag{4.2.0.21}$$

and this estimated mode used for W^*, this Gaussian-work bias becomes

$$B(M) = -\beta^{-1} \ln \left\{ \frac{1}{2} \operatorname{erfc} \left[-\frac{1}{\sqrt{2}} \left(\sqrt{\mathbf{W}_L \left[\frac{1}{2\pi} (M-1)^2 \right]} - \beta\sigma \right) \right] \right\}, \tag{4.2.0.22}$$

and depends solely on

$$\Pi = \sqrt{\mathbf{W}_L \left[\frac{1}{2\pi} (M-1)^2 \right]} - \beta\sigma. \tag{4.2.0.23}$$

By using Eq. 4.2.0.17, it can be rewritten as

$$\Pi = \sqrt{\mathbf{W}_L \left[\frac{1}{2\pi} (M-1)^2 \right]} - \sqrt{2\beta(\bar{W} - \Delta A)}. \tag{4.2.0.24}$$

Wu and Kofke suggested that the number of work samples (non-equilibrium trajectories) should be sufficient to ensure $\Pi > 0.5$.[64] This idea has also been applied to the estimate of the bias in bridge estimators such as BAR.[65]

4.3 Kullback–Leibler divergence

The convergence rate of thermodynamic perturbation heavily depends on the similarity between the simulated Hamiltonian (A) and the target one (B). Kullback–Leibler divergence offers a quantitative way to measure the similarity by

$$s_A = \int_\Gamma \mathrm{d}\mathbf{x}\, p_A(\mathbf{x}) \ln \left[\frac{p_A(\mathbf{x})}{p_B(\mathbf{x})} \right]. \tag{4.3.0.1}$$

s_A is also known as the relative entropy in Information theory. Similarly, we can also have

$$s_B = \int_\Gamma \mathrm{d}\mathbf{x}\, p_B(\mathbf{x}) \ln \left[\frac{p_B(\mathbf{x})}{p_A(\mathbf{x})} \right]. \tag{4.3.0.2}$$

With the Boltzmann statistics,

$$p_A(\mathbf{x}) = e^{-\beta U(\mathbf{x})}/Q_A. \tag{4.3.0.3}$$

The relative entropies can be written as

$$\begin{aligned} s_A &= \int_\Gamma \mathrm{d}\mathbf{x}\, \frac{e^{-\beta U_A(\mathbf{x})}}{Q_A} \ln \left[\frac{e^{-\beta U_A(\mathbf{x})}}{Q_A} \frac{Q_B}{e^{-\beta U_B(\mathbf{x})}} \right] \\ &= \int_\Gamma \mathrm{d}\mathbf{x}\, \frac{e^{-\beta U_A(\mathbf{x})}}{Q_A} \left\{ \beta \left[U_B(\mathbf{x}) - U_A(\mathbf{x}) \right] - \beta\Delta A \right\} && (4.3.0.4) \\ &= \beta \left\langle \Delta U \right\rangle_A - \beta\Delta A && (4.3.0.5) \end{aligned}$$

and

$$s_B = -\beta \left\langle \Delta U \right\rangle_B + \beta\Delta A, \tag{4.3.0.6}$$

which are exactly the dissipated work[66].

Appendix A

Statistical Uncertainty in the Estimator for Correlated Time Series Data

"It is not the estimate or the forecast that matters so much as the degree of confidence with the opinion."

– Nassim Nicholas Taleb

Suppose we have a time series of correlated sequential observations of the randomly sampled variable X denoted as $\{x_n\}, n = 1, \ldots, N$ that come from a stationary, time-reversible stochastic process. The expectation of X can be estimated as the time average of the samples

$$\hat{X} = \frac{1}{N} \sum_{n=1}^{N} x_n. \tag{A.0.0.1}$$

Because of the existence of correlation among the samples, the variance for the expectation, which is defined as

$$\delta^2 \hat{X} \equiv \left\langle \left(\hat{X} - \left\langle \hat{X} \right\rangle \right)^2 \right\rangle = \left\langle \hat{X}^2 \right\rangle - \left\langle \hat{X} \right\rangle^2, \tag{A.0.0.2}$$

is complicated. We first take Eq. A.0.0.1 into Eq. A.0.0.2, and split the sum into one term capturing the variance in the observations and a remaining term capturing the correlation between the observations as

$$\begin{aligned} \delta^2 \hat{X} =& \frac{1}{N^2} \sum_{n,n'=1}^{N} \left[\langle x_n x_{n'} \rangle - \langle x_n \rangle \langle x_{n'} \rangle \right] \\ =& \frac{1}{N^2} \sum_{n=1}^{N} \left[\left\langle x_n^2 \right\rangle - \langle x_n \rangle^2 \right] + \frac{1}{N^2} \sum_{n \neq n'=1}^{N} \left[\langle x_n x_{n'} \rangle - \langle x_n \rangle \langle x_{n'} \rangle \right] \end{aligned} \tag{A.0.0.3}$$

Because of the stationarity, it becomes

$$\begin{aligned}\delta^2\hat{X} =& \frac{1}{N}\left[\left\langle x_n^2\right\rangle - \langle x_n\rangle^2\right] \\ &+ \frac{1}{N^2}\sum_{t=1}^{N-1}(N-t)\left[\langle x_n x_{n+t}\rangle + \langle x_{n+t}x_n\rangle - \langle x_n\rangle\langle x_{n+t}\rangle - \langle x_{n+t}\rangle\langle x_n\rangle\right]\end{aligned} \tag{A.0.0.4}$$

and because of the time-reversibility, it can be further simplified to

$$\begin{aligned}\delta^2\hat{X} =& \frac{1}{N}\left[\left\langle x_n^2\right\rangle - \langle x_n\rangle^2\right] \\ &+ \frac{2}{N}\sum_{t=1}^{N-1}\left(\frac{N-t}{N}\right)\left[\langle x_n x_{n+t}\rangle - \langle x_n\rangle\langle x_{n+t}\rangle\right] \\ \equiv& \frac{{\sigma_x}^2}{N}(1+2\tau) = \frac{{\sigma_x}^2}{N/g},\end{aligned} \tag{A.0.0.5}$$

where ${\sigma_x}^2$, statistical inefficiency g, and autocorrelation time τ (in units of the sampling interval) are given by

$${\sigma_x}^2 \equiv \left\langle x_n^2\right\rangle - \langle x_n\rangle^2 \tag{A.0.0.6}$$

$$\tau \equiv \sum_{t=1}^{N-1}\left(\frac{N-t}{N}\right)C_t \tag{A.0.0.7}$$

$$C_t = \frac{\langle x_n x_{n+t}\rangle - \langle x_n\rangle\langle x_n\rangle}{{\sigma_x}^2} \tag{A.0.0.8}$$

$$g \equiv 1 + 2\tau \tag{A.0.0.9}$$

The quantity $g \equiv 1 + 2\tau > 1$ can be regarded as a statistical inefficiency, in that N/g gives the effective number of *uncorrelated* configurations contained in the time series.

Appendix B

The Optimal Mean of Independent Measurements with Uncertainties

Suppose we have N measurements of a quantity x, which are denoted as $\{x_i\}$, with i = 1, ..., N. Each measurement has a variance $\delta^2 x_i$. To find the optimal mean of this data set, we first write the mean of $\{x_i\}$ as a weighted average of them

$$\bar{x} = \sum_{i=1}^{N} a_i x_i, \tag{B.0.0.1}$$

in which a_i are the normalized weights, i.e.

$$\sum_{i=1}^{N} a_i = 1. \tag{B.0.0.2}$$

According to the error propagation rule, if the measurements are independent, the variance of the mean $\bar{x}$ can be written as

$$\delta^2 \bar{x} = \sum_{i=1}^{N} {a_i}^2 \delta^2 x_i. \tag{B.0.0.3}$$

Minimizing $\delta^2 \bar{x}$ with respect to a_i under the constraint of Eq. B.0.0.2 using the Lagrange multiplier λ, we find

$$\begin{aligned} \frac{\partial L}{\partial a_j} =& \frac{\partial}{\partial a_j} \left[\sum_{i=1}^{N} {a_i}^2 \delta^2 x_i + \lambda \left(1 - \sum_{i=1}^{N} a_i \right) \right] \\ =& 2 a_j \delta^2 x_j - \lambda \\ =& 0 \end{aligned} \tag{B.0.0.4}$$

for all x_j. It can be easily identified that a_j is inversely proportional to $\delta^2 x_j$, i.e.

$$a_j = \frac{\left(\delta^2 x_j\right)^{-1}}{\sum\limits_{i=1}^{N} \left(\delta^2 x_i\right)^{-1}}, \tag{B.0.0.5}$$

and

$$\bar{x} = \sum_{j=1}^{N} \frac{\left(\delta^2 x_j\right)^{-1}}{\sum\limits_{i=1}^{N} \left(\delta^2 x_i\right)^{-1}} x_j, \tag{B.0.0.6}$$

with

$$\delta^2 \bar{x} = \sum_{j=1}^{N} \frac{\left(\delta^2 x_j\right)^{-1}}{\left(\sum\limits_{i=1}^{N} \left(\delta^2 x_i\right)^{-1}\right)^2} = \frac{1}{\sum\limits_{i=1}^{N} \left(\delta^2 x_i\right)^{-1}}. \tag{B.0.0.7}$$

Appendix C

The Relationship between the ΔU Distributions in Forward and Backward TP

In a forward TP calculation between H_0 and H_1, the distribution of ΔU is

$$f(\Delta U) = \frac{\int e^{-\beta H_0}\delta(H_1 - H_0 - \Delta U)\,\mathrm{d}x}{\int e^{-\beta H_0}\,\mathrm{d}x}. \tag{C.0.0.1}$$

While in a backward TP, the distribution is

$$\begin{aligned}
g(\Delta U) =& \frac{\int e^{-\beta H_1}\delta(H_1 - H_0 - \Delta U)\,\mathrm{d}x}{\int e^{-\beta H_1}\,\mathrm{d}x} \\
=& \frac{f(\Delta U)\int e^{-\beta H_1}\delta(H_1 - H_0 - \Delta U)\,\mathrm{d}x \int e^{-\beta H_0}\,\mathrm{d}x}{\int e^{-\beta H_1}\,\mathrm{d}x \int e^{-\beta H_0}\delta(H_1 - H_0 - \Delta U)\,\mathrm{d}x} \\
=& f(\Delta U)\frac{\int e^{-\beta H_0}\,\mathrm{d}x}{\int e^{-\beta H_1}\,\mathrm{d}x}\frac{\int e^{-\beta H_1}\delta(H_1 - H_0 - \Delta U)\,\mathrm{d}x}{\int e^{-\beta H_0}\delta(H_1 - H_0 - \Delta U)\,\mathrm{d}x} \\
=& f(\Delta U)e^{\beta\Delta A}\frac{\int e^{-\beta\Delta H}e^{-\beta H_0}\delta(H_1 - H_0 - \Delta U)\,\mathrm{d}x}{\int e^{-\beta H_0}\delta(H_1 - H_0 - \Delta U)\,\mathrm{d}x} \\
=& f(\Delta U)e^{\beta\Delta A}\frac{e^{-\beta\Delta U}\int e^{-\beta H_0}\delta(H_1 - H_0 - \Delta U)\,\mathrm{d}x}{\int e^{-\beta H_0}\delta(H_1 - H_0 - \Delta U)\,\mathrm{d}x} \\
=& f(\Delta U)e^{\beta\Delta A}e^{-\beta\Delta U}
\end{aligned} \tag{C.0.0.2}$$

Or it can be written as

$$g(\Delta U)e^{\beta\Delta U} = f(\Delta U)e^{\beta\Delta A}, \tag{C.0.0.3}$$

which was first shown by Shing and Gubbins for Widom insertion and deletion[67].

By looking at this equation and the integral in energy space for TP (Eq. 3.1.1.15),

$$\Delta A = -\frac{1}{\beta} \ln \int \exp(-\beta \Delta U) f(\Delta U)\, \mathrm{d}\Delta U, \tag{C.0.0.4}$$

it can be easily realized that the integrand is proportional to the probability of energy difference sampled from state 1 ($g(\Delta U)$) and the free energy difference ΔA can be estimated reliably only when the sampling at state 0 covers the representative configurations of state 1.

Appendix D

Cumulant Expansion for the Free Energy Difference in Thermodynamic Perturbation Calculations

In Thermodynamic Perturbation, the free energy difference between states are expressed as

$$\Delta A = -\beta^{-1} \ln \langle \exp\left[-\beta \Delta U\right] \rangle_0. \tag{D.0.0.1}$$

Due to the exponential term on the right-hand-side of this equation, the convergence is usually slow. This expression can be expanded into Taylor series, and the leading terms can converge much faster than the complete sum.

Using the Taylor expansion for $e^x = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} x^n$ first, we have

$$\exp\left[-\beta \Delta U\right] = 1 + (-\beta \Delta U) + \frac{1}{2!}(-\beta \delta U)^2 + \frac{1}{3!}(-\beta \Delta U)^3 + \cdots . \tag{D.0.0.2}$$

Then Eq. D.0.0.1 becomes

$$\Delta A = -\beta^{-1} \ln \left(1 - \beta \left\langle \Delta U \right\rangle_0 + \frac{1}{2}\beta^2 \left\langle \Delta U^2 \right\rangle_0 - \frac{1}{6}\beta^3 \left\langle \Delta U^3 \right\rangle_0 + \cdots \right). \tag{D.0.0.3}$$

Using the Taylor expansion for $\ln(1+x) = x - \frac{1}{2}x^2 + \frac{2}{3!}x^3 - \frac{6}{4!}x^4 + \cdots$, we

have

$$\begin{aligned}
\Delta A = & \langle \Delta U \rangle_0 - \frac{1}{2}\beta \left\langle \Delta U^2 \right\rangle_0 + \frac{1}{6}\beta^2 \left\langle \Delta U^3 \right\rangle_0 + \cdots \\
& + \frac{1}{2}\left(\beta \langle \Delta U \rangle_0^2 - \beta^2 \langle \Delta U \rangle_0 \left\langle \Delta U^2 \right\rangle_0 + \frac{1}{4}\beta^3 \left\langle \Delta U^2 \right\rangle_0^2 + \cdots \right) \\
& - \frac{1}{3}\left(-\beta^2 \langle \Delta U \rangle_0^3 + \cdots \right) \\
& + \cdots \\
= & \langle \Delta U \rangle_0 - \frac{\beta}{2}\left(\left\langle \Delta U^2 \right\rangle_0 - \langle \Delta U \rangle_0^2 \right) \\
& + \frac{\beta^2}{6}\left(\left\langle \Delta U^3 \right\rangle_0 - 3 \left\langle \Delta U^2 \right\rangle_0 \langle \Delta U \rangle_0 + 2 \langle \Delta U \rangle_0^3 \right) + \cdots \qquad \text{(D.0.0.4)}
\end{aligned}$$

Appendix E

MBAR Returns to BAR When Only Two States Are Considered

When there are only two states, the free energy in Eq. 3.1.5.7 for the 1st state in MBAR becomes

$$\begin{aligned} f_1 =& -\beta_1{}^{-1} \ln \sum_{n=1}^{N} \frac{\exp\left(-\beta_1 U_1(\mathbf{R}_n)\right)}{\sum_{k=1}^{2} N_k \exp\left(\beta_k f_k - \beta_k U_k(\mathbf{R}_n)\right)} \\ =& -\beta_1{}^{-1} \ln \sum_{j=1}^{2} \sum_{n=1}^{N_j} \frac{\exp\left(-\beta_1 U_1(\mathbf{R}_{jn})\right)}{\sum_{k=1}^{2} N_k \exp\left(\beta_k f_k - \beta_k U_k(\mathbf{R}_{jn})\right)}, \end{aligned} \tag{E.0.0.1}$$

or equivalently we have

$$1 = \sum_{n=1}^{N} \frac{\exp\left(\beta_1 f_1 - \beta_1 U_1(\mathbf{R}_n)\right)}{N_1 \exp\left(\beta_1 f_1 - \beta_1 U_1(\mathbf{R}_n)\right) + N_2 \exp\left(\beta_2 f_2 - \beta_2 U_2(\mathbf{R}_n)\right)}, \tag{E.0.0.2}$$

$$\begin{aligned} N_1 =& \sum_{n=1}^{N_1} \frac{1}{1 + \frac{N_2}{N_1} \exp\left(\Delta f - \Delta U(\mathbf{R}_{1n})\right)} \\ &+ \sum_{n=1}^{N_2} \frac{1}{1 + \frac{N_2}{N_1} \exp\left(\Delta f - \Delta U(\mathbf{R}_{2n})\right)} \end{aligned} \tag{E.0.0.3}$$

where $\Delta f = \beta_2 f_2 - \beta_1 f_1$ and $\Delta U = \beta_2 U_2 - \beta_1 U_1$. We further define $M = -\ln \frac{N_2}{N_1}$, then

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

$$\sum_{n=1}^{N_1} \frac{1}{1+\exp\left(-\Delta f + \Delta U(\mathbf{R}_{1n}) + M\right)} = \sum_{n=1}^{N_2} \frac{1}{1+\exp\left(\Delta f - \Delta U(\mathbf{R}_{2n}) - M\right)}$$

$$\sum_{n=1}^{N_1} f\left(-\Delta f + \Delta U(\mathbf{R}_{1n}) + M\right) = \sum_{n=1}^{N_2} f\left(\Delta f - \Delta U(\mathbf{R}_{2n}) - M\right)$$

$$N_1 \left\langle f\left(-\Delta f + \Delta U(\mathbf{R}_{1n}) + M\right)\right\rangle_1 = N_2 \left\langle f\left(\Delta f - \Delta U(\mathbf{R}_{2n}) - M\right)\right\rangle_2$$

$$\frac{\left\langle f\left(\Delta f - \Delta U(\mathbf{R}_{2n}) - M\right)\right\rangle_2}{\left\langle f\left(-\Delta f + \Delta U(\mathbf{R}_{1n}) + M\right)\right\rangle_1} = \frac{N_1}{N_2},$$

which is Eq. 3.1.3.12.

Appendix F

MBAR is a binless form of WHAM

Maybe you have already noticed that MBAR and WHAM have very similar forms for the free energy. So you may want to ask if there is any connection between MBAR and WHAM. The answer is YES. MBAR is a binless form of WHAM.[50] Let us follow Zhang et al[68] and rewrite Eq. 3.1.4.21 into an integral form

$$f_i = -\ln \int \Omega \exp\left(-\beta_i U\right) \mathrm{d}U. \tag{F.0.0.1}$$

Taking Eq. 3.1.4.20 into Eq. F.0.0.1, we find

$$f_i = -\ln \int \frac{\sum_{k=1}^{K} H_k(U) \exp\left(-\beta_i U\right)}{\sum_{k=1}^{K} N_k \exp\left(f_k - \beta_k U\right)} \mathrm{d}U, \tag{F.0.0.2}$$

where ${g_{mk}}^{-1}$ has been omitted and H_{mk} has been changed to continuous form $H_k(U)$. From the definition,

$$H_k(U) = \sum_{\mathbf{R}}^{(k)} \delta(U(\mathbf{R}) - U). \tag{F.0.0.3}$$

Taking Eq. F.0.0.3 into Eq. F.0.0.2, we have

$$\begin{aligned} f_i &= -\ln \sum_{k=1}^{K} \sum_{\mathbf{R}}^{(k)} \frac{\exp\left(-\beta_i U(\mathbf{R})\right)}{\sum_{k=1}^{K} N_k \exp\left(f_k - \beta_k U(\mathbf{R})\right)} \\ &= -\ln \sum_{n=1}^{N} \frac{\exp\left(-\beta_i U_i(\mathbf{R_n})\right)}{\sum_{k=1}^{K} N_k \exp\left(f_k - \beta_k U_k(\mathbf{R_n})\right)}, \end{aligned} \tag{F.0.0.4}$$

which is Eq. 3.1.5.7.

Bibliography

[1] Niels Hansen and Wilfred F. van Gunsteren. Practical Aspects of Free-Energy Calculations: A Review. *J. Chem. Theory Comput.*, 10(7):2632–2647, 2014.

[2] Yuqing Deng and Benoît Roux. Calculation of Standard Binding Free Energies: Aromatic Molecules in the T4 Lysozyme L99A Mutant. *J. Chem. Theory Comput.*, 2(5):1255–1273, 2006.

[3] Hyung-June Woo and Benoît Roux. Calculation of Absolute Protein-ligand Binding Free Energy from Computer Simulations. *Proc. Natl. Acad. Sci. U. S. A.*, 102(19):6825–6830, 2005.

[4] Yuqing Deng and Benoît Roux. Computations of Standard Binding Free Energies with Molecular Dynamics Simulations. *J. Phys. Chem. B*, 113(8):2234–2246, 2009.

[5] James C. Gumbart, Benoît Roux, and Christophe Chipot. Standard Binding Free Energies from Computer Simulations: What Is the Best Strategy? *J. Chem. Theory Comput.*, 9(1):794–802, 2013.

[6] Ulrich H. E. Hansmann and Yuko Okamoto. Prediction of Peptide Conformation by Multicanonical Algorithm: New Approach to the Multiple-Minima Problem. *J. Comput. Chem.*, 14(11):1333–1338, 1993.

[7] Ulrich H.E. Hansmann. Parallel Tempering Algorithm for Conformational Studies of Biological Molecules. *Chem. Phys. Lett.*, 281(1):140–150, 1997.

[8] Yuji Sugita and Yuko Okamoto. Replica-exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.*, 314(1):141–151, 1999.

[9] Yilin Meng and Adrian E. Roitberg. Constant pH Replica Exchange Molecular Dynamics in Biomolecules Using a Discrete Protonation Model. *J. Chem. Theory Comput.*, 6(4):1401–1412, 2010.

[10] Juyong Lee, Benjamin T. Miller, Ana Damjanović, and Bernard R. Brooks. Constant pH Molecular Dynamics in Explicit Solvent with En-

veloping Distribution Sampling and Hamiltonian Exchange. *J. Chem. Theory Comput.*, 10(7):2738–2750, 2014.

[11] E. Marinari and G. Parisi. Simulated tempering: A new monte carlo scheme. *Europhys. Lett.*, 19(6):451–458, 1992.

[12] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov. New approach to monte carlo calculation of the free energy: Method of expanded ensembles. *J. Chem. Phys.*, 96(3):1776–1783, 1992.

[13] Charles J. Geyer and Elizabeth A. Thompson. Annealing markov chain monte carlo with applications to ancestral inference. *J. Am. Stat. Assoc.*, 90(431):909–920, 1995.

[14] G. M. Torrie and J. P. Valleau. Nonphysical Sampling Distributions in Monte Carlo Free-energy Estimation: Umbrella Sampling. *J. Comput. Phys.*, 23(2):187–199, 1977.

[15] Giovanni Ciccotti, Raymond Kapral, and Eric Vanden-Eijnden. Blue Moon Sampling, Vectorial Reaction Coordinates, and Unbiased Constrained Dynamics. *ChemPhysChem*, 6(9):1809–1814, 2005.

[16] E. A. Carter, Giovanni Ciccotti, James T. Hynes, and Raymond Kapral. Constrained Reaction Coordinate Dynamics for the Simulation of Rare Events. *Chem. Phys. Lett.*, 156(5):472–477, 1989.

[17] Eric Darve and Andrew Pohorille. Calculating Free Energies Using Average Force. *J. Chem. Phys.*, 115(20):9169–9183, 2001.

[18] Eric Darve, David Rodríguez-Gómez, and Andrew Pohorille. Adaptive Biasing Force Method for Scalar and Vector Free Energy Calculations. *J. Chem. Phys.*, 128(14):144120, 2008.

[19] Jerry B. Abrams, Lula Rosso, and Mark E. Tuckerman. Efficient and Precise Solvation Free Energies via Alchemical Adiabatic Molecular Dynamics. *J. Chem. Phys.*, 125(7):074115, 2006.

[20] Pan Wu, Xiangqian Hu, and Weitao Yang. λ-Metadynamics Approach To Compute Absolute Solvation Free Energy. *J. Phys. Chem. Lett.*, 2(17):2099–2103, 2011.

[21] Fugao Wang and D. P. Landau. Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States. *Phys. Rev. Lett.*, 86:2050–2053, 2001.

[22] Yves F. Atchadé and Jun S. Liu. The Wang-Landau Algorithm in General State Spaces: Applications and Convergence Analysis. *Stat. Sinica*, 20(1):209–233, 2010.

[23] Faming Liang, Chuanhai Liu, and Raymond J. Carroll. Stochastic Approximation in Monte Carlo Computation. *J. Am. Stat. Assoc.*, 102(477):305–320, 2007.

[24] Alessandro Laio and Michele Parrinello. Escaping Free-energy Minima. *Proc. Natl. Acad. Sci. U. S. A.*, 99(20):12562–12566, 2002.

[25] Alessandro Barducci, Giovanni Bussi, and Michele Parrinello. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.*, 100:020603, 2008.

[26] Omar Valsson, Pratyush Tiwary, and Michele Parrinello. Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint. *Annu. Rev. Phys. Chem.*, 67(1):159–184, 2016.

[27] James F. Dama, Michele Parrinello, and Gregory A. Voth. Well-Tempered Metadynamics Converges Asymptotically. *Phys. Rev. Lett.*, 112:240602, 2014.

[28] Omar Valsson and Michele Parrinello. Variational Approach to Enhanced Sampling and Free Energy Calculations. *Phys. Rev. Lett.*, 113:090601, 2014.

[29] Lianqing Zheng, Mengen Chen, and Wei Yang. Random Walk in Orthogonal Space to Achieve Efficient Free-energy Simulation of Complex Systems. *Proc. Natl. Acad. Sci. U. S. A.*, 105(51):20227–20232, 2008.

[30] David A. Pearlman and Peter A. Kollman. The Lag between the Hamiltonian and the System Configuration in Free Energy Perturbation Calculations. *J. Chem. Phys.*, 91(12):7831–7839, 1989.

[31] Xianjun Kong and Charles L. Brooks III. λ-dynamics: A New Approach to Free Energy Calculations. *J. Chem. Phys.*, 105(6):2414–2423, 1996.

[32] Clara D. Christ and Wilfred F. van Gunsteren. Enveloping Distribution Sampling: A Method to Calculate Free Energy Differences from a Single Simulation. *J. Chem. Phys.*, 126(18):184110, 2007.

[33] Clara D. Christ and Wilfred F. van Gunsteren. Simple, Efficient, and Reliable Computation of Multiple Free Energy Differences from a Single Simulation: A Reference Hamiltonian Parameter Update Scheme for Enveloping Distribution Sampling (EDS). *J. Chem. Theory Comput.*, 5(2):276–286, 2009.

[34] Weinan E, Weiqing Ren, and Eric Vanden-Eijnden. String Method for the Study of Rare Events. *Phys. Rev. B*, 66:052301, 2002.

[35] Weinan E, Weiqing Ren, and Eric Vanden-Eijnden. Simplified and Improved String Method for Computing the Minimum Energy Paths in Barrier-Crossing Events. *J. Chem. Phys.*, 126(16):164103, 2007.

[36] Eric Vanden-Eijnden and Maddalena Venturoli. Revisiting the Finite Temperature String Method for the Calculation of Reaction Tubes and Free Energies. *J. Chem. Phys.*, 130(19):194103, 2009.

[37] Robert W. Zwanzig. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.*, 22(8):1420, 1954.

[38] William L. Jorgensen and Laura L. Thomas. Perspective on Free-Energy Perturbation Calculations for Chemical Equilibria. *J. Chem. Theory Comput.*, 4(6):869–876, 2008.

[39] Christopher Jarzynski. Rare Events and the Convergence of Exponentially Averaged Work Values. *Phys. Rev. E*, 73:046105, 2006.

[40] John G. Kirkwood. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.*, 3(5):300–313, 1935.

[41] Himanshu Paliwal and Michael R. Shirts. A Benchmark Test Set for Alchemical Free Energy Transformations and Its Use to Quantify Error in Common Free Energy Methods. *J. Chem. Theory Comput.*, 7(12):4115–4134, 2011.

[42] Charles H. Bennett. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys.*, 22(2):245–268, 1976.

[43] Gavin E. Crooks. Path-ensemble Averages in Systems Driven Far From Equilibrium. *Phys. Rev. E*, 61:2361–2366, 2000.

[44] Michael R. Shirts, Eric Bair, Giles Hooker, and Vijay S. Pande. Equilibrium Free Energies from Nonequilibrium Measurements Using Maximum-Likelihood Methods. *Phys. Rev. Lett.*, 91:140601, 2003.

[45] Herman J. C. Berendsen. *A Student's Guide to Data and Error Analysis*. Cambridge University Press, Cambridge, 2011.

[46] Alan M. Ferrenberg and Robert H. Swendsen. Optimized Monte Carlo Data Analysis. *Phys. Rev. Lett.*, 63:1195–1198, 1989.

[47] John D. Chodera, William C. Swope, Jed W. Pitera, Chaok Seok, and Ken A. Dill. Use of the Weighted Histogram Analysis Method for the Analysis of Simulated and Parallel Tempering Simulations. *J. Chem. Theory Comput.*, 3(1):26–41, 2007.

[48] Marc Souaille and Benoît Roux. Extension to the Weighted Histogram Analysis Method: Combining Umbrella Sampling with Free Energy Calculations. *Comput. Phys. Commun.*, 135:40–57, 2001.

[49] Emilio Gallicchio, Michael Andrec, Anthony K. Felts, and Ronald M. Levy. Temperature Weighted Histogram Analysis Method, Replica Exchange, and Transition Paths. *J. Phys. Chem. B*, 109(14):6722–6731, 2005.

[50] Zhiqiang Tan, Emilio Gallicchio, Mauro Lapelosa, and Ronald M. Levy. Theory of Binless Multi-state Free Energy Estimation with Applications to Protein-ligand Binding. *J. Chem. Phys.*, 136(14):144102, 2012.

[51] Michael R. Shirts and John D. Chodera. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.*, 129(12):124105, 2008.

[52] Michael R. Shirts. Reweighting from the Mixture Distribution as a Better Way to Describe the Multistate Bennett Acceptance Ratio. arXiv.org, 2017. https://arxiv.org/abs/1704.00891.

[53] Xinqiang Ding, Jonah Z. Vilseck, and Charles L. Brooks. Fast Solver for Large Scale Multistate Bennett Acceptance Ratio Equations. *J. Chem. Theory Comput.*, 15(2):799–802, 2019.

[54] Johannes Kästner and Walter Thiel. Bridging the Gap between Thermodynamic Integration and Umbrella Sampling Provides a Novel Analysis Method: "Umbrella Integration". *J. Chem. Phys.*, 123(14):144104, 2005.

[55] Johannes Kästner and Walter Thiel. Analysis of the Statistical Error in Umbrella Sampling Simulations by Umbrella Integration. *J. Chem. Phys.*, 124(23):234106, 2006.

[56] Johannes Kästner. Umbrella Integration with Higher-Order Correction Terms. *J. Chem. Phys.*, 136(23):234102, 2012.

[57] Christopher Jarzynski. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.*, 78:2690, 1997.

[58] Gavin E. Crooks. Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems. *J. Statis. Phys.*, 90:1481, 1998.

[59] Gerhard Hummer and Attila Szabo. Free Energy Reconstruction from Nonequilibrium Single-molecule Pulling Experiments. *Proc. Natl. Acad. Sci. U. S. A.*, 98(7):3658–3661, 2001.

[60] Hao Wu, Antonia S. J. S. Mey, Edina Rosta, and Frank Noé. Statistically Optimal Analysis of State-discretized Trajectory Data from Multiple Thermodynamic States. *J. Chem. Phys.*, 141(21):214106, 2014.

[61] Jessica M.J. Swanson, Richard H. Henchman, and J. Andrew McCammon. Revisiting Free Energy Calculations: A Theoretical Connection to MM/PBSA and Direct Calculation of the Association Free Energy. *Biophys. J.*, 86(1):67–74, 2004.

[62] Pavel V. Klimovich, Michael R. Shirts, and David L. Mobley. Guidelines for the Analysis of Free Energy Calculations. *J. Comput. Aided Mol. Des.*, 29(5):397–411, 2015.

[63] Nandou Lu and David A. Kofke. Accuracy of Free-Energy Perturbation Calculations in Molecular Simulation. I. Modeling. *J. Chem. Phys.*, 114(17):7303–7311, 2001.

[64] Di Wu and David A. Kofke. Model for Small-Sample Bias of Free-Energy Calculations Applied to Gaussian-Distributed Nonequilibrium Work Measurements. *J. Chem. Phys.*, 121(18):8742–8747, 2004.

[65] Brian K. Radak. Finite-Sample Bias in Free Energy Bridge Estimators. *J. Chem. Phys.*, 151(3):034105, 2019.

[66] Di Wu and David A. Kofke. Phase-Space Overlap Measures. I. Fail-Safe Bias Detection in Free Energies Calculated by Molecular Simulation. *J. Chem. Phys.*, 123(5):054103, 2005.

[67] K. S. Shing and K. E. Gubbins. The Chemical Potential in Dense Fluids and Fluid Mixtures via Computer Simulation. *Mol. Phys.*, 46(5):1109–1128, 1982.

[68] Cheng Zhang, Chun-Liang Lai, and B. Montgomery Pettitt. Accelerating the Weighted Histogram Analysis Method by Direct Inversion in the Iterative Subspace. *Mol. Simulat.*, 42(13):1079–1089, 2016.