Enhanced Sampling
&
Free Energy Calculations

Methods Based on Equilibrium and Non-equilibrium Simulations

Ye Mei [1]
Pengfei Li [2]

July 24, 2018

[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], Ms. Meiting Wang, Ms. Wei Liu and Ms. 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 many difficulties in three aspects, i.e. Hamiltonians, sampling efficiency and 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, `http://pubs.acs.org/toc/jctcce/10/7`).

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 state as

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

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 systems denoted by 0 and 1, respectively

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

In most situations 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. Or they may be characterized by different values of a macroscopic parameter, such as temperature. Finally, they may correspond to different regions in the phase space accessible to the system

$$Q_0 = \frac{1}{N!h^{3N}} \int_{\Gamma_0} \exp\left[-\beta H(\mathbf{x}, \mathbf{p}_x)\right] \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x \tag{1.0.0.9}$$

$$Q_1 = \frac{1}{N!h^{3N}} \int_{\Gamma_1} \exp\left[-\beta H(\mathbf{x}, \mathbf{p}_x)\right] \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}_x \tag{1.0.0.10}$$

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.

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

$$\rho(x) = \frac{1}{Z} \exp\left(-\beta E(x)\right), \tag{1.0.0.11}$$

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

$$\langle O \rangle = \frac{\int O(\mathbf{x}) \exp\left(-\beta E(\mathbf{x})\right) \mathrm{d}\mathbf{x}}{Z}. \tag{1.0.0.12}$$

Besides state free energy, sometimes we are also interested in the free energy profile along one or several degrees of freedom

$$\begin{aligned} A(\xi) =& -\beta^{-1} \ln Z(\xi) \\ =& -\beta^{-1} \ln \int \exp\left(-\beta U\right) \delta(\xi - \xi(\mathbf{x}))\, \mathrm{d}\mathbf{x} \\ =& -\beta^{-1} \ln \int \exp\left(-\beta U(\xi(\mathbf{x}) = \xi)\right) |J|\, \mathrm{d}q_1 \cdots \mathrm{d}q_{N-1}, \end{aligned} \tag{1.0.0.13}$$

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

$$\begin{aligned} \frac{\partial A(\xi)}{\partial \xi} =& -\beta^{-1} \frac{\int \frac{\partial}{\partial \xi} \left(e^{-\beta U} |J|\right) \mathrm{d}q_1 \cdots \mathrm{d}q_{N-1}}{\int e^{-\beta U} \delta(\xi - \xi(\mathbf{x}))\, \mathrm{d}\mathbf{x}} \\ =& \frac{\int e^{-\beta U} \left[\frac{\partial U}{\partial \xi} - \beta^{-1} \frac{1}{|J|} \frac{\partial |J|}{\partial \xi}\right] |J|\, \mathrm{d}q_1 \cdots \mathrm{d}q_{N-1}}{\int e^{-\beta U} \delta(\xi - \xi(\mathbf{x}))\, \mathrm{d}\mathbf{x}} \\ =& \frac{\int e^{-\beta U} \left[\frac{\partial U}{\partial \xi} - \beta^{-1} \frac{\partial \ln |J|}{\partial \xi}\right] |J|\, \mathrm{d}q_1 \cdots \mathrm{d}q_{N-1}}{\int e^{-\beta U} \delta(\xi - \xi(\mathbf{x}))\, \mathrm{d}\mathbf{x}} \\ =& \frac{\int e^{-\beta U} \left[\frac{\partial U}{\partial \xi} - \beta^{-1} \frac{\partial \ln |J|}{\partial \xi}\right] \delta(\xi - \xi(\mathbf{x}))\, \mathrm{d}\mathbf{x}}{\int e^{-\beta U} \delta(\xi - \xi(\mathbf{x}))\, \mathrm{d}\mathbf{x}} \\ =& \left\langle \frac{\partial U}{\partial \xi} - \beta^{-1} \frac{\partial \ln |J|}{\partial \xi} \right\rangle_{\xi}. \end{aligned} \tag{1.0.0.14}$$

Here, $-\frac{\partial U}{\partial \xi}+\beta^{-1}\frac{\partial \ln |J|}{\partial \xi}$ is the generalized force on ξ 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.*

In order to illustrate how the simulations and free energy methods to be discussed below are used in real problems, let us take protein-ligand binding

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

as an example. The equilibrium constant, k_b, is defined as

$$K_b = \frac{[\text{P-L}]}{[\text{P}][\text{L}]}, \tag{1.0.0.15}$$

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 $G_{bind} \equiv -\beta^{-1} \ln \left[K_b C^0\right]$, 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

$$K_b = \frac{1}{[L]} \frac{N \int_{site} \mathrm{d}(\mathbf{1}) \int_{bulk} \mathrm{d}(\mathbf{2}) \cdots \int_{bulk} \mathrm{d}(\mathbf{N}) \int \mathrm{d}\mathbf{X} e^{-\beta U}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \int_{bulk} \mathrm{d}(\mathbf{2}) \cdots \int_{bulk} \mathrm{d}(\mathbf{N}) \int \mathrm{d}\mathbf{X} e^{-\beta U}}, \tag{1.0.0.16}$$

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

$$\begin{aligned} K_b =& \frac{1}{[L]} \frac{N \int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta U}}{V_{bulk} \int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r_1} - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta U}} \\ =& \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta U}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r_1} - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta U}}. \end{aligned} \tag{1.0.0.17}$$

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

$$\begin{aligned} K_b = \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta U}}{Z_1} \times \frac{Z_1}{Z_2} \times \cdots \times \frac{Z_{n-1}}{Z_n} \times \\ \frac{Z_n}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r_1} - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta U}} \end{aligned} \tag{1.0.0.18}$$

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, 5]. For each ratio in Eq. 1.0.0.18, 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 1.0.0.18 is now realized as

$$\begin{aligned} K_b = & \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta U_1}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c]}} \times \\ & \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c]}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c+U_t]}} \times \\ & \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c+U_t]}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c+U_t+U_r]}} \times \\ & \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c+U_t+U_r]}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_0+U_c+U_t+U_r]}} \times \\ & \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_0+U_c+U_t+U_r]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_0+U_c+U_t]}} \times \\ & \frac{\int_{bulk} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U_0+U_c+U_t]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U_0+U_c]}} \times \\ & \frac{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U_0+U_c]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c]}} \times \\ & \frac{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U_1+U_c]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathbf{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta U_1}}. \end{aligned} \tag{1.0.0.19}$$

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. Equa-

tion 1.0.0.18 is now realized as

$$
\begin{aligned}
K_b = & \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta U}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c]}} \times \\
& \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c]}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c+U_o]}} \times \\
& \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c+U_o]}}{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c+U_o+U_a]}} \times \\
& \frac{\int_{site} \mathrm{d}(\mathbf{1}) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c+U_o+U_a]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathrm{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c+U_o]}} \times \\
& \frac{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathrm{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c+U_o]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathrm{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c]}} \times \\
& \frac{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathrm{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta[U+U_c]}}{\int_{bulk} \mathrm{d}(\mathbf{1}) \delta(\mathrm{r}_1 - \mathbf{r}^*) \int \mathrm{d}\mathbf{X} e^{-\beta U}}. \qquad (1.0.0.20)
\end{aligned}
$$

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 low 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 $\exp[-\beta E(\mathbf{R})]$, where $\beta = 1/kT$ is called the inverse temperature. k 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 kT, which is 6 kcal/mol at room temperature (300 K). If the barrier is 1.5 kcal/mol higher, it takes the system about 1 μs (10 times longer) in average to go over the barrier. If the barrier height reaches 9 kcal/mol, it takes 10 μs. And so on. With modern computers, the longest all-atom molecular dynamics simulation for biological systems is probably 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 more than 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 Sugita and Okamoto in 1999[6] 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.

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

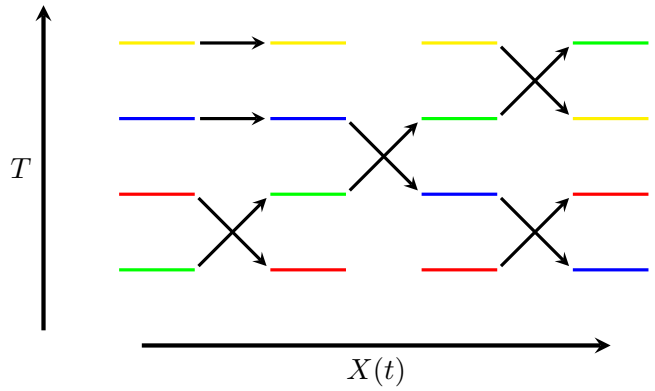

Figure 2.1: A schematic representation of replica exchange molecular dynamics.

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

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

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

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

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

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

$$\Delta G_A = \Delta G_A^0 + \beta^{-1} \ln \frac{[A]}{C_0},$$

in which ΔG_0 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

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

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

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

in which $[A]\,[B] \,/\, [C]\, C_0$ 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

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

with

$$K_a = \frac{[P_A]\,[H^+]}{[P_{AH^+}]\,C_0}$$

The fraction of the deprotonated species is calculated as

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

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

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

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

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

Generally, the absolute value of ΔG_0 is hardly computable. A relative protonation free energy $\Delta\Delta G$ is perferred and is more reliable. Theoretically, the reference state can be any state you like. But the protonation free energy of the dipeptide at pK_a is often used. The reference protonation process can be written as

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

The reference free energy change is

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

where R_H is the coordinates of the specific H atom and the other degrees-of-freedom (DoF) are denoted as R_O. 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,

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

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

$$\begin{aligned}\Delta\Delta G_0 = &- \beta^{-1} \ln \frac{\int \exp\left(-\beta E^{ele}_{P_AH^+}\right) \exp\left(-\beta E^{other}_{P_AH^+}\right) \mathrm{d}R_O}{\int \exp\left(-\beta E^{ele}_{AH^+}\right) \exp\left(-\beta E^{other}_{AH^+}\right) \mathrm{d}R_O} \\ &\cdot \frac{\int \exp\left(-\beta E_A\right) \mathrm{d}R_O}{\int \exp\left(-\beta E_{P_A}\right) \mathrm{d}R_O}. \end{aligned} \tag{2.1.2.8}$$

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

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

$$\cdot \frac{\int \exp\left(-\beta E_A\right) \mathrm{d}R_O}{\int \exp\left(-\beta E^{ele}_{AH^+}\right) \exp\left(-\beta E_A\right) \mathrm{d}R_O} \tag{2.1.2.10}$$

$$\begin{aligned} =&- \beta^{-1} \ln \left\langle \exp\left(-\beta E^{ele}_{P_AH^+}\right) \right\rangle_{P_A} \\ &+ \beta^{-1} \ln \left\langle \exp\left(-\beta E^{ele}_{AH^+}\right) \right\rangle_{A} \\ =&\Delta G^{ele}_{P_AH^+} - \Delta G^{ele}_{AH^+} \end{aligned} \tag{2.1.2.11}$$

Therefore,

$$-\beta^{-1} \ln 10 \cdot pK_a = \Delta G^{ele}_{P_AH^+} - \Delta G^{ele}_{AH^+} - \beta^{-1} \ln 10 \cdot pK_a^{ref}.$$

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

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

In the above equation, $\Delta G^{ele}_{AH^+}$ can be obtained from a free energy calculation of the model system by alchemically annihilation of the proton. However, $\Delta G^{ele}_{P_AH^+}$ is unknown. Approximately, it can be replaced with $\Delta H^{ele}_{P_AH^+}$ averaged over a few snapshots.[7] In order to accelerate the convergence, this pH-REMD is often coupled with other enhanced simulation methods, such as T-REMD[7] and EDS-REMD[8] (see section 2.6).

2.2 Umbrella Sampling

Umbrella Sampling method was proposed by Torrie and Valleau in 1977,[9] and is still widely used nowadays. Suppose we are studying a transition process between states such as conversion between two dominant conformations or a chemical reaction, and these two states are separated by a high barrier relative to kT. Therefore, the transition is a rare event. A schematic representation of the free energy landscape is shown in Fig. 2.2.

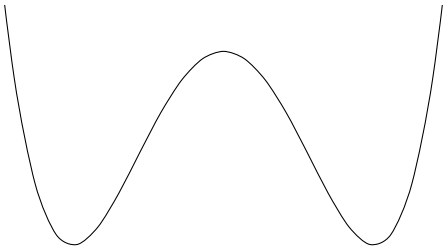

Figure 2.2: A typical free energy surface. Two free energy wells are separated by a barrier higher than kT.

Sometimes, we are interested in not only these two dominate states but also the states in between. 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 it can be 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 brutal 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. And the simulations are performed on these potential energy surfaces

$$U_i(\mathbf{R}) = U_0(\mathbf{R}) + \Delta U_i(\xi). \tag{2.2.0.1}$$

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 whole region is well sampled. Ensemble average under U_0 can be calculated

from the ensembles generated under the biased Hamiltonians U via

$$\begin{aligned}\langle X(\mathbf{R})\rangle_0 =& \frac{\int X(\mathbf{R}) \exp\left[-\beta U_0(\mathbf{R})\right] \mathrm{d}\mathbf{R}}{\int \exp\left[-\beta U_0(\mathbf{R})\right] \mathrm{d}\mathbf{R}} \\ =& \frac{\int X(\mathbf{R}) \exp\left[\beta \Delta U_i(\mathbf{R})\right] \exp\left[-\beta U_i(\mathbf{R})\right] \mathrm{d}\mathbf{R}}{\int \exp\left[\beta \Delta U_i(\mathbf{R})\right] \exp\left[-\beta U_i(\mathbf{R})\right] \mathrm{d}\mathbf{R}} \\ =& \frac{\langle X \exp\left(\beta \Delta U_i\right)\rangle_i}{\langle \exp\left(\beta \Delta U_i\right)\rangle_i}. \end{aligned} \tag{2.2.0.2}$$

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).

2.3 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

$$H(\mathbf{R}, \lambda) = H_{Rxn}(\mathbf{R}, \lambda) + \frac{m_\lambda}{2}\dot{\lambda}^2 + U^*(\lambda), \tag{2.3.0.1}$$

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.[10] λ-dynamics can also be coupled with metadynamics,[11] which will be introduced in Sec. 2.4.

2.4 Metadynamics

Metadynamics, vividly called flooding method, was first suggested by Laio and Parrinello in 2002.[12] 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 was 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. At that moment, you got the exact shape of the original free energy landscape.

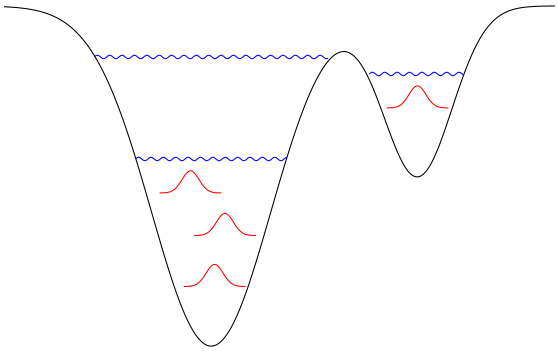

Figure 2.3: A schematic representation of metadynamics. The free energy well is gradually filled up with small Gaussians, and a transition is facilitated.

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

$$V_G(\mathbf{S},t) = \int_0^t \mathrm{d}t' \omega \exp\left(-\sum_{i=1}^{d} \frac{[\mathbf{S}_i(R) - \mathbf{S}_i(R(t'))]^2}{2\sigma_i^2}\right) \qquad (2.4.0.1)$$

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,

$$V_G(\mathbf{S}, t \to \infty) = -G(\mathbf{S}) + C. \tag{2.4.0.2}$$

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.

2.5 Orthogonal Space Random Walk

The orthogonal space random walk (OSRW) was developed by Yang in 2008.[13] 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

$$\Delta G(\xi_i \to \xi_f) = \int_{\xi_i}^{\xi_f} \left.\frac{\partial G}{\partial \xi}\right|_{\xi'} \mathrm{d}\xi' = \int_{\xi_i}^{\xi_f} \left\langle \frac{\partial U}{\partial \xi} - \beta^{-1} \frac{\partial \ln |J|}{\partial \xi} \right\rangle_{\xi'} \mathrm{d}\xi', \quad (2.5.0.1)$$

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 long 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.[14] 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. 2.4), is added to the Hamiltonian of the system recursively, which has a functional form like

$$h \exp\left(-\frac{|\xi - \xi(t_i)|^2}{2{w_1}^2}\right) \exp\left(-\frac{|F_\xi - F_\xi(t_i)|^2}{2{w_2}^2}\right). \quad (2.5.0.2)$$

The overall biasing potential

$$G(\xi, F_\xi) = \sum_{t_i} h \exp\left(-\frac{|\xi - \xi(t_i)|^2}{2{w_1}^2}\right) \exp\left(-\frac{|F_\xi - F_\xi(t_i)|^2}{2{w_2}^2}\right). \quad (2.5.0.3)$$

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.[15] 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

$$\left.\frac{\partial G}{\partial \xi}\right|_{\xi'} = \langle F_\xi \rangle_{\xi'} = \frac{\sum F_\xi \exp\left[\beta G(\xi, F_\xi)\right]\delta(\xi - \xi')}{\sum \exp\left[\beta G(\xi, F_\xi)\right]\delta(\xi - \xi')}, \tag{2.5.0.4}$$

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

$$\Delta G(\xi) = \int_{\xi_i}^{\xi} \left.\frac{\partial G}{\partial \xi}\right|_{\xi'} \mathrm{d}\xi'. \tag{2.5.0.5}$$

2.6 Enveloping Distribution Sampling

Enveloping distribution sampling method was first proposed by Christ and van Gunsteren in 2007.[16] When calculating the free energy difference between states A and B,

$$\Delta G_{BA} = G_B - G_A = -\beta^{-1} \ln \frac{Q_B}{Q_A}, \tag{2.6.0.1}$$

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.

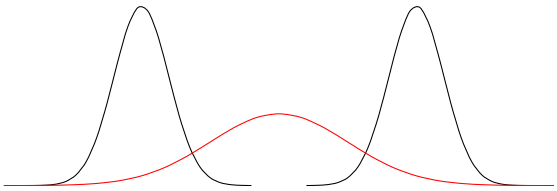

Figure 2.4: The configuration distributions under two Hamiltonians have no visible overlap as shown by solid black curves. A reference state (shown as the red curve) that has remarkable overlap with both states can be introduced to accelerate the convergence of the free energy calculations using, for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

$$\Delta G_{BA} = \Delta G_{BR} - \Delta G_{AR} = -\beta^{-1} \ln \frac{\langle e^{-\beta(H_B - H_R)} \rangle_R}{\langle e^{-\beta(H_A - H_R)} \rangle_R}, \tag{2.6.0.2}$$

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

$$H_R(\mathbf{r}) = -(s\beta)^{-1} \ln \left(e^{-s\beta H_A(\mathbf{r})} + e^{-s\beta H_B(\mathbf{r})} \right), \tag{2.6.0.3}$$

where s is a scale factor that modulates the mixing[17] 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. 2.6.0.3, 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.

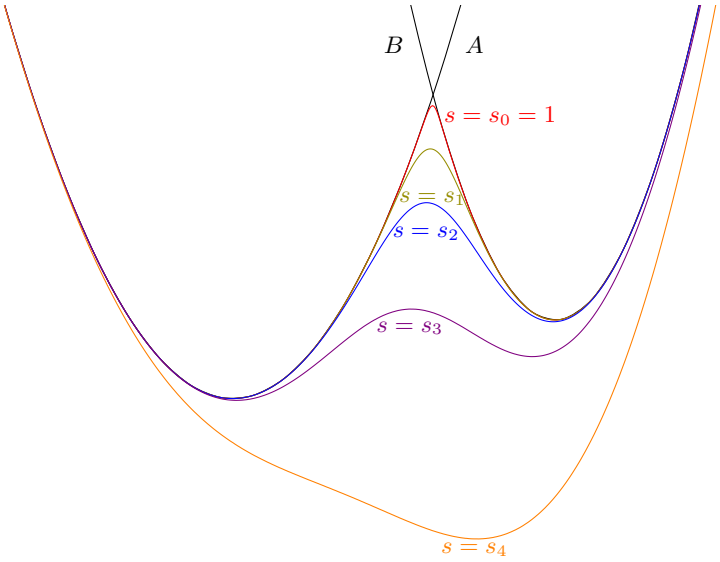

Figure 2.5: State A and state B have only negligible overlap at high energy regions. The reference state generated by the mixing of state A and state B is characterized by s. Increasing s may lower the barrier between the dominant wells.

The force is also a mixing quantity from two Hamiltonians as

$$\mathbf{F}_R^i = -\frac{\partial H_R}{\partial \mathbf{r}^i} = \frac{e^{-s\beta H_A(\mathbf{r})}}{e^{-s\beta H_A(\mathbf{r})} + e^{-s\beta H_B(\mathbf{r})}} \left(-\frac{\partial H_A(\mathbf{r})}{\partial \mathbf{r}^i}\right) + \frac{e^{-s\beta H_B(\mathbf{r})}}{e^{-s\beta H_A(\mathbf{r})} + e^{-s\beta H_B(\mathbf{r})}} \left(-\frac{\partial H_B(\mathbf{r})}{\partial \mathbf{r}^i}\right). \quad (2.6.0.4)$$

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,[18] and Landau and Lifshitz, independently, and probably by Peierls[19].

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 connected 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

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

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

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

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

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

where ΔU is the difference in the potential energy between the target and the reference states. The integration implied by the statistical average is now carried out over particle coordinates only.

If we reverse the reference and the target systems, and repeat the same derivation, using the same convention for ΔA and ΔU as before, we obtain

$$\Delta A = \frac{1}{\beta}\ln\left\langle\exp(\beta\Delta U)\right\rangle_1, \tag{3.1.1.10}$$

Although expressions Eq. 3.1.1.9 and Eq. 3.1.1.10 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. This means that the reference system

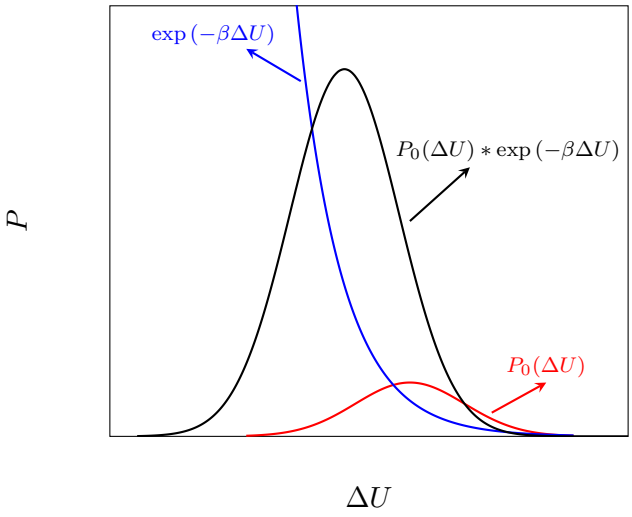

Figure 3.1: $P_0(\Delta U)$, the Boltzmann factor $\exp(-\beta\Delta U)$ and their product, which is the integrand in Eq. 3.1.1.11. The low-ΔU tail of the integrand is poorly sampled with $P_0(\Delta U)$ and, therefore, is known with low statistical accuracy. However, it provides an important contribution to the integral.

should be that with higher entropy, and the transformation should proceed in the direction in which the entropy decreases.

Equation 3.1.1.9 and Eq. 3.1.1.10, are formally exact for any perturbation. However, this does not mean that they can always be successfully applied. 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

$$\Delta A = -\frac{1}{\beta} \int \exp(-\beta\Delta U) P_0(\Delta U)\, \mathrm{d}\Delta U, \tag{3.1.1.11}$$

If U_0 and U_1 were the 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 the ideal Gaussian case, but still has a "Gaussian-like" shape. This indicates that the value of the integral in Eq. 3.1.1.11 depends on the low-energy tail of the distribution.

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

$$P_0(\Delta U) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\Delta U - \langle \Delta U \rangle_0)^2}{2\sigma^2}\right] \tag{3.1.1.12}$$

where

$$\sigma^2 = \left\langle \Delta U^2 \right\rangle_0 - \langle \Delta U \rangle_0^2 \tag{3.1.1.13}$$

to Eq. 3.1.1.11, we obtain

$$\exp(-\beta \Delta A) = \frac{C}{\sqrt{2\pi}\sigma} \int \exp\left[-\frac{(\Delta U - \langle \Delta U \rangle_0 - \beta\sigma^2)^2}{2\sigma^2} \right] \mathrm{d}\Delta U \tag{3.1.1.14}$$

Here, C is independent of ΔU

$$C = \exp\left[-\beta(\langle \Delta U \rangle_0 - \frac{1}{2}\beta\sigma^2) \right] \tag{3.1.1.15}$$

If $P_0(\Delta U)$ is Gaussian, the integral in Eq. 3.1.1.14 can be evaluated analytically using cumulant expansion

$$\Delta A = \langle \Delta U \rangle_0 - \frac{1}{2}\beta\sigma^2. \tag{3.1.1.16}$$

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

$$\Delta A = \langle \Delta U \rangle_0 - \frac{1}{2}\beta\sigma^2 + \frac{\beta^2}{6}\left(\left\langle \Delta U^3 \right\rangle_0 - 3\left\langle \Delta U^2 \right\rangle_0 \langle \Delta U \rangle_0 + 2\langle \Delta U \rangle_0^3 \right). \tag{3.1.1.17}$$

The overlap matrix[20] is a useful metric in characterizing the magnitude of overlap in the phase space. It is recommended to be used as a consistency check, as in the case of the TP-based methods. If the weight of each of the N samples x_n (collected from all K states) in the ith state as:

$$W_{n,i}(x_n) = \frac{e^{\beta G_i - \beta U_i(x_n)}}{\sum_{k=1}^{K} N_k e^{\beta G_k - \beta U_k(x_n)}} \tag{3.1.1.18}$$

Then, the overlap matrix is a $K \times K$ matrix with entries:

$$O_{i,j} = \sum_{n=1}^{N} \frac{N_i e^{\beta G_i - \beta U_i(x_n)}}{\sum_{k=1}^{K} N_k e^{\beta G_k - \beta U_k(x_n)}} \frac{e^{\beta G_j - \beta U_j(x_n)}}{\sum_{l=1}^{L} N_l e^{\beta G_l - \beta U_l(x_n)}} \tag{3.1.1.19}$$

3.1.2 Thermodynamic Integration

Thermodynamic Integration (TI) method was originally proposed by Kirkwood.[21].

If the free energy, A, is a continuous function of λ then we can write:

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

Now,

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

Thus,

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

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 can write the following for $\partial Q(\lambda)/\partial\lambda$:

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

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

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

Thus, the basic TI formula is

$$\Delta A = \int_{\lambda=0}^{\lambda=1}\left\langle \frac{\partial H(\mathbf{x},\mathbf{p}_x,\lambda)}{\partial \lambda}\right\rangle_\lambda \mathrm{d}\lambda \tag{3.1.2.7}$$

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 molecule dynamic 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

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

The exact calculation of ΔA requires an infinite number of ensemble average for λ ranging from 0 to 1. Therefore, the integral in Eq. 3.1.2.7 needs to be approximated, e.g., by a summation over a number of discrete points λ_i, which leads to

$$\Delta A = \sum_i \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda_i} \Delta\lambda_i. \tag{3.1.2.9}$$

A finite number of λ_i values between 0 and 1 are chosen and for each of them a complete molecule dynamic 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.

The accuracy of TI integral formula depends on the exact method for the numerical integration.[22] In addition to summation method, the simplest numerical integration is to evaluate the integrand at the midpoint:

$$\Delta A \simeq \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda=\frac{1}{2}} \tag{3.1.2.10}$$

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

3.1.3 Bennett Acceptance Ratio

Bennett acceptance ratio was developed by Bennett in 1976,[23] and was re-discovered by Crooks[24] and Shirts et al[25] 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

$$\frac{Q_0}{Q_1} = \frac{< M(U_0 - U_1) >_1}{< M(U_1 - U_0) >_0}. \tag{3.1.3.5}$$

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 ($covar(x_1, x_2) = 0$),[26]

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

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

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

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

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

To minimize it with respect to W, we have

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

Substituting this into Eq. 3.1.3.6 yields

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

where

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

and f denotes the Fermi function

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

3.1.4 Weighted Histogram Analysis Method

The weighted histogram analysis method is a generalization of the histogram method developed by Ferrenberg and Swendsen.[27]

Weighted Histogram Analysis Method for Parallel Tempering

The following derivation quite follows Ref. [28]. The central quantity in statistical mechanics is partition function Z, which in textbook is often written as

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

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

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

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 fallen 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)

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

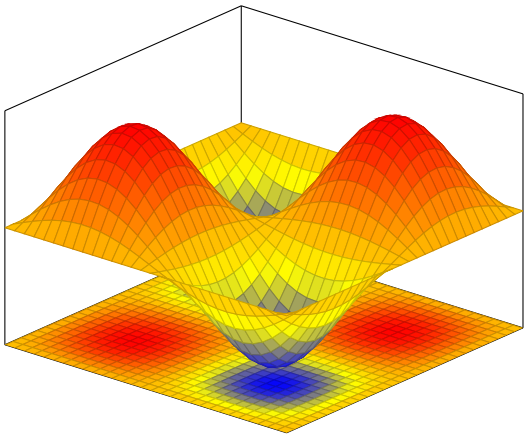

Figure 3.2: A sample histogram in 2D space, for instance potential energy and a reaction coordinate ξ.

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

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

which is the weighted average of density of states of all the trajectories with the weight reversely proportional to the uncertainties (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. $H_{mk} \ll N_k$. The expectation of H_{mk} can be related to the optimum estimator of the density of states, i.e.

$$\overline{H_{mk}} = N_k \Delta U \Omega_m \exp\left(f_k - \beta_k U_m\right). \tag{3.1.4.16}$$

Then we have

$$\delta^2 H_{mk} = g_{mk} N_k \Delta U \Omega_m \exp\left(f_k - \beta_k U_m\right) \tag{3.1.4.17}$$

and

$$\delta^2 \Omega_{mk} = \frac{\Omega_m}{{g_{mk}}^{-1} N_k \Delta U \exp\left(f_k - \beta_k U_m\right)}. \tag{3.1.4.18}$$

Taking Eq. 3.1.4.13 and Eq. 3.1.4.18 into Eq. 3.1.4.15, we find

$$\Omega_m = \frac{\sum\limits_{k=1}^{K} {g_{mk}}^{-1} H_{mk}}{\sum\limits_{k=1}^{K} {g_{mk}}^{-1} N_k \Delta U \exp\left(f_k - \beta_k U_m\right)}, \tag{3.1.4.19}$$

in which

$$f_k = -\ln \sum_{m=1}^{M} \Omega_m \Delta U \exp\left(-\beta_k U_m\right). \tag{3.1.4.20}$$

Obviously, Eq. 3.1.4.19 and Eq. 3.1.4.20 must be solved iteratively. Applying the error propagation rule to Eq. 3.1.4.19 and using Eq. 3.1.4.17, the uncertainty of Ω_m is given by

$$\delta^2 \Omega_m = \frac{\Omega_m}{\sum\limits_{k=1}^{K} {g_{mk}}^{-1} N_k \Delta U \exp\left(f_k - \beta_k U_m\right)}, \tag{3.1.4.21}$$

and the relative uncertainty is given by

$$\frac{\delta^2\Omega_m}{\Omega_m^2} = \left[\sum_{k=1}^{K} {g_{mk}}^{-1} H_{mk}\right]^{-1}. \tag{3.1.4.22}$$

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.[29] 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. A initial guess of the values $f_i^{(b)0}$ is used on the right side of Eq. 3.1.4.38 to compute a new set of values $f_i^{(b)1}$, which is in turn used as the new guess to compute $f_i^{(b)1}$ and so on until convergence is reached. Then, the unbiased free energy corresponding to the histogram can be calculated by,

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

To get rid of the constant C in the Eq. 3.1.4.36, 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. [30], 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,

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

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. 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

$$c_{k,l} = \exp\left[-\left(\beta_k - \beta_0\right) E_l\right], \tag{3.1.4.42}$$

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

$$c_{k,l} = \exp\left(-\beta_0 w_{k,l}\right). \tag{3.1.4.43}$$

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

$$\begin{aligned} &P(n_{k,1}, n_{k,2}, \ldots, n_{k,L} | p_{k,1}, p_{k,2}, \ldots, p_{k,L}) = \\ &\frac{\left(\sum_l n_{k,l}\right)!}{\prod_l n_{k,l}!} \prod_{l=1}^{L} \left(p_{k,l}\right)^{n_{k,l}} \propto \prod_{l=1}^{L} \left(f_k c_{k,l} p_l^0\right)^{n_{k,l}}. \end{aligned} \tag{3.1.4.44}$$

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

$$\begin{aligned} P(n_{1,1}, \ldots, n_{1,L}; \ldots; n_{K,1}, \ldots, n_{K,L} | p_1^0, \ldots, p_L^0) \propto \\ \prod_{k=1}^{K} \prod_{l=1}^{L} \left(f_k c_{k,l} p_l^0\right)^{n_{k,l}}, \end{aligned} \tag{3.1.4.45}$$

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. 3.1.4.45 with respect to $p_1^0, \ldots, p_L^0$ and are given by solutions of the simultaneous nonlinear equations

$$p_l^0 = \frac{\sum_{k=1}^{K} n_{k,l}}{\sum_{k=1}^{K} N_k f_k c_{k,l}} \text{ (for each } l\text{)} \tag{3.1.4.46}$$

and

$$f_k = \left\{ \sum_l c_{k,l} p_l^0 \right\}^{-1}, \tag{3.1.4.47}$$

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

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.[31] The following derivation quite follows Ref. [32] 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. $p_k(\mathbf{R}_n)$ 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[9], 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,

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

3.1.6 Non-Equilibrium Work

Non-Equilibrium Work (NEW) method for equilibrium free energy calculations was proposed by Jarzynski.[33]. In 1997, Jarzynski showed

$$\langle \exp\left[-\beta W(\tau)\right]\rangle = \exp\left(-\beta \Delta A\right), \tag{3.1.6.1}$$

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. 3.1.6.1 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.[34]

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

$$\Delta A = \lim_{\tau \to \infty} \int_0^\tau \left. \frac{\partial H\left[\mathbf{x}(t); \lambda\right]}{\partial \lambda} \right|_{\lambda = \lambda(t)} \dot{\lambda}(t)\, \mathrm{d}t, \tag{3.1.6.2}$$

where $\dot{\lambda}(t)$ is the time derivative of the coupling parameter λ. In Eq. 3.1.6.2, the limit of $\tau \to \infty$ ensures that the transformation is performed infinitely slowly, and thus reversibly. The right-hand side of Eq. 3.1.6.2 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

$$\langle W(\tau)\rangle \geq \Delta A, \tag{3.1.6.3}$$

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

$$W(\tau) = \int_0^\tau \frac{\partial H[\mathbf{x}(t);\lambda]}{\partial \lambda}\bigg|_{\lambda=\lambda(t)} \dot{\lambda}(t)\,\mathrm{d}t \tag{3.1.6.4}$$

The equality in Eq. 3.1.6.3 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. 3.1.6.1, immediately leads to the second law in the form of Eq. 3.1.6.3 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. 3.1.6.2 is the formula of TI

$$\Delta A = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial H(\mathbf{x}, \mathbf{p}_x, \lambda)}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda \tag{3.1.6.5}$$

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,

$$\lim_{\tau\to 0} W(\tau) = H(\mathbf{x}(0); \lambda = 1) - H(\mathbf{x}(0); \lambda = 0) \tag{3.1.6.6}$$

Substituting the Eq. 3.1.6.6 into the Eq. 3.1.6.1, the formula of TP can be recovered

$$\Delta A = -\frac{1}{\beta} \ln \langle \exp[-\beta \Delta H(\mathbf{x}, \mathbf{p}_x)]\rangle_0 \,, \tag{3.1.6.7}$$

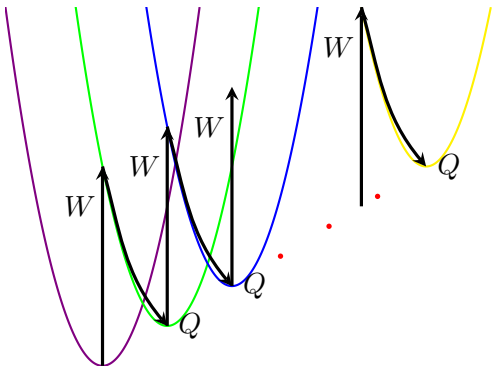

Figure 3.3: The accumulation of work and heat along a nonequilibrium trajectory. The work is defined as the energy change when the coupling parameter switches from λ_i to λ_{i+1} with the coordinates fixed, while the dissipated heat is defined as the energy relaxation when the coordinate change with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

In Ref. [34], 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

$$\frac{p_f[w = W(\tau)]}{p_b[w = -\underline{W}(\tau)]} = \exp[\beta(w - \Delta A)], \tag{3.1.6.8}$$

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. 3.1.6.1 follows from Eq. 3.1.6.8 simply by integration over w because the probability densities are normalized to 1:

$$\int p_f(W)e^{-\beta W}\,\mathrm{d}W = \int p_b(W)e^{-\beta\Delta A}\,\mathrm{d}W, \tag{3.1.6.9}$$

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

Following the Crooks Fluctuation Theorem (CFT),[34] 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.[25] In this approach, ΔA is calculated via

$$\sum_{i=1}^{n_F} \frac{1}{1 + \exp\left[\beta(M + W_i - \Delta A)\right]} = \sum_{j=1}^{n_R} \frac{1}{1 + \exp\left[-\beta(M + W_j - \Delta A)\right]}, \tag{3.1.6.10}$$

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. [25].

3.1.7 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.[35]. 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 $\gamma_i^{(k)}$

$$\pi_i^{(k)} = f^{(k)} \pi_i \gamma_i^{(k)}, \tag{3.1.7.1}$$

$$f^{(k)} = \frac{1}{\sum_l \pi_l \gamma_l^{(k)}}, \tag{3.1.7.2}$$

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

$$L_{\mathrm{WHAM}} = \prod_k \prod_i (\pi_i^{(k)})^{N_i^{(k)}} \tag{3.1.7.3}$$

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 $\boldsymbol{\pi}$

$$L_{\mathrm{MSM}} = \prod_i \prod_j p_{ij}^{c_{ij}}, \tag{3.1.7.4}$$

$$s.t. \quad \pi_i p_{ij} = \pi_j p_{ji} \quad \text{for all } i, j, \tag{3.1.7.5}$$

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. 3.1.7.1 and 3.1.7.2. The resulting likelihood of all $\mathbf{P}^{(k)}$ and $\boldsymbol{\pi}$, based on simulations at all thermodynamic states can be formulated as

$$L_{\mathrm{TRAM}} = \prod_k \prod_i \prod_j (p_{ij}^{(k)})^{c_{ij}^{(k)}}, \tag{3.1.7.6}$$

$$s.t. \quad \pi_i^{(k)} p_{ij}^{(k)} = \pi_j^{(k)} p_{ji}^{(k)} \quad \text{for all } i, j, k. \tag{3.1.7.7}$$

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. 3.1.7.6 and 3.1.7.7 difficult, and it can neither be achieved by WHAM, nor by existing MSM estimators.*

Taking natural logarithm on the TRAM likelihood, we find

$$\ln L_{\mathrm{TRAM}} = \sum_{k=1}^{K} \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}^{(k)} \ln p_{ij}^{(k)}, \tag{3.1.7.8}$$

with constraints

$$\pi_i \gamma_i^{(k)} p_{ij}^{(k)} = \pi_j \gamma_j^{(k)} p_{ji}^{(k)} \quad \text{for all } i, j, k, \tag{3.1.7.9}$$

which is from Eqs. 3.1.7.1 and 3.1.7.7 with $f^{(k)}$ canceled. In addition, $\mathbf{P}^{(k)}$ and $\boldsymbol{\pi}$ should satisfy the normalization conditions

$$\sum_j p_{ij}^{(k)} = 1 \quad \forall i, k, \tag{3.1.7.10}$$

$$\sum_j \pi_j = 1. \tag{3.1.7.11}$$

The normalization of $\boldsymbol{\pi}^{(k)}$ is naturally satisfied due to Eqs. 3.1.7.1 and 3.1.7.2. 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

$$\sum_k \sum_j \frac{(c_{ij}^{(k)} + c_{ji}^{(k)})\gamma_i^{(k)}\pi_i\nu_j^{(k)}}{\gamma_i^{(k)}\pi_i\nu_j^{(k)} + \gamma_j^{(k)}\pi_j\nu_i^{(k)}} = \sum_k \sum_j c_{ji}^{(k)}, \tag{3.1.7.12}$$

$$\sum_j \frac{(c_{ij}^{(k)} + c_{ji}^{(k)})\gamma_j^{(k)}\pi_j}{\gamma_i^{(k)}\pi_i\nu_j^{(k)} + \gamma_j^{(k)}\pi_j\nu_i^{(k)}} = 1 \tag{3.1.7.13}$$

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

$$\pi_i^{init} := 1/n, v_i^{(k),init} := \sum_j c_{ij}^{(k)}, \tag{3.1.7.14}$$

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

$$v_i^{(k),new} := v_i^{(k)} \sum_j \frac{(c_{ij}^{(k)} + c_{ji}^{(k)})\gamma_i^{(k)}\pi_j}{\gamma_i^{(k)}\pi_i\nu_j^{(k)} + \gamma_j^{(k)}\pi_j\nu_i^{(k)}}, \tag{3.1.7.15}$$

$$\pi_i^{new} := \frac{\sum_{k,j} c_{ji}^{(k)}}{\sum_{k,j} \frac{(c_{ij}^{(k)} + c_{ji}^{(k)})\gamma_i^{(k)}\nu_j^{(k)}}{\gamma_i^{(k)}\pi_i\nu_j^{(k)} + \gamma_j^{(k)}\pi_j\nu_i^{(k)}}}. \tag{3.1.7.16}$$

3.2 Approximate Methods

3.2.1 Molecular Mechanics/Poisson-Boltzmann Surface Area

The following derivation follows Ref. [36]. 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

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

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

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

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

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

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

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

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

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

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.

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 statistical uncertainty 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 Data Set

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 inversed proportional to $\delta^2 x_j$, i.e.

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

and

$$\bar{x} = \sum_{i=1}^{N} \frac{\delta^2 {x_j}^{-1}}{\sum_{i=1}^{N} \delta^2 {x_j}^{-1}} x_i, \tag{B.0.0.6}$$

Appendix C

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

$$\begin{aligned} f_1 =& -{\beta_1}^{-1} \ln \sum_{n=1}^{N} \frac{\exp\left(-\beta_1 U_1(\mathbf{R}_n)\right)}{\sum_{k=1}^{2} N_k \exp\left(\beta_k f_k - \beta_k U_k(\mathbf{R}_n)\right)} \\ =& -{\beta_1}^{-1} \ln \sum_{j=1}^{2} \sum_{n=1}^{N_j} \frac{\exp\left(-\beta_1 U_1(\mathbf{R}_{jn})\right)}{\sum_{k=1}^{2} N_k \exp\left(\beta_k f_k - \beta_k U_k(\mathbf{R}_{jn})\right)}, \end{aligned} \tag{C.0.0.1}$$

or equivalently we have

$$1 = \sum_{n=1}^{N} \frac{\exp\left(\beta_1 f_1 - \beta_1 U_1(\mathbf{R}_n)\right)}{N_1 \exp\left(\beta_1 f_1 - \beta_1 U_1(\mathbf{R}_n)\right) + N_2 \exp\left(\beta_2 f_2 - \beta_2 U_2(\mathbf{R}_n)\right)}, \tag{C.0.0.2}$$

$$\begin{aligned} N_1 =& \sum_{n=1}^{N_1} \frac{1}{1 + \frac{N_2}{N_1} \exp\left(\Delta f - \Delta U(\mathbf{R}_{1n})\right)} \\ &+ \sum_{n=1}^{N_2} \frac{1}{1 + \frac{N_2}{N_1} \exp\left(\Delta f - \Delta U(\mathbf{R}_{2n})\right)} \end{aligned} \tag{C.0.0.3}$$

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 D

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.[37] Let us follow Zhang et al[38] and rewrite Eq. 3.1.4.20 into an integral form

$$f_i = -\ln \int \Omega \exp\left(-\beta_i U\right) \mathrm{d}U. \tag{D.0.0.1}$$

Taking Eq. 3.1.4.19 into Eq. D.0.0.1, we find

$$f_i = -\ln \int \frac{\sum_{k=1}^{K} H_k(U) \exp\left(-\beta_i U\right)}{\sum_{k=1}^{K} N_k \exp\left(f_k - \beta_k U\right)} \mathrm{d}U, \tag{D.0.0.2}$$

where ${g_{mk}}^{-1}$ has been omitted and H_{mk} has been changed to continuous form $H_k(U)$. From the definition,

$$H_k(U) = \sum_{\mathbf{R}}^{(k)} \delta(U(\mathbf{R}) - U). \tag{D.0.0.3}$$

Taking Eq. D.0.0.3 into Eq. D.0.0.2, we have

$$\begin{aligned} f_i &= -\ln \sum_{k=1}^{K} \sum_{\mathbf{R}}^{(k)} \frac{\exp\left(-\beta_i U(\mathbf{R})\right)}{\sum_{k=1}^{K} N_k \exp\left(f_k - \beta_k U(\mathbf{R})\right)} \\ &= -\ln \sum_{n=1}^{N} \frac{\exp\left(-\beta_i U_i(\mathbf{R_n})\right)}{\sum_{k=1}^{K} N_k \exp\left(f_k - \beta_k U_k(\mathbf{R_n})\right)}, \end{aligned} \tag{D.0.0.4}$$

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] Yuji Sugita and Yuko Okamoto. Replica-exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.*, 314(1):141–151, 1999.

[7] 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.

[8] Juyong Lee, Benjamin T. Miller, Ana Damjanović, and Bernard R. Brooks. Constant pH Molecular Dynamics in Explicit Solvent with Enveloping Distribution Sampling and Hamiltonian Exchange. *J. Chem. Theory Comput.*, 10(7):2738–2750, 2014.

[9] 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.

[10] 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.

[11] 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.

[12] Alessandro Laio and Michele Parrinello. Escaping Free-energy Minima. *Proc. Natl. Acad. Sci. U. S. A.*, 99(20):12562–12566, 2002.

[13] 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.

[14] 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.

[15] Xianjun Kong and Charles L. Brooks III. λ-dynamics: A New Approach to Free Energy Calculations. *J. Chem. Phys.*, 105(6):2414–2423, 1996.

[16] 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.

[17] 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.

[18] Robert W. Zwanzig. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.*, 22(8):1420, 1954.

[19] 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.

[20] Pavel V. Klimovich, Michael R. Shirts, and David L. Mobley. Guidelines for the Analysis of Free Energy Calculations. *Journal of Computer-Aided Molecular Design*, 29(5):397–411, 2015.

[21] John G. Kirkwood. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.*, 3(5):300–313, 1935.

[22] 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. *Journal of Chemical Theory and Computation*, 7(12):4115–4134, 2011.

[23] Charles H. Bennett. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys.*, 22(2):245–268, 1976.

[24] Gavin E. Crooks. Path-ensemble Averages in Systems Driven Far From Equilibrium. *Phys. Rev. E*, 61:2361–2366, 2000.

[25] 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.

[26] Herman J. C. Berendsen. *A Student's Guide to Data and Error Analysis*. Cambridge University Press, Cambridge, 2011.

[27] Alan M. Ferrenberg and Robert H. Swendsen. Optimized Monte Carlo data analysis. *Phys. Rev. Lett.*, 63:1195–1198, 1989.

[28] 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.

[29] M Souaille and B Roux. Extension to the Weighted Histogram Analysis Method: Combining Umbrella Sampling with Free Energy Calculations. *Comput. Phys. Commun.*, 135:40–57, 2001.

[30] 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.

[31] Michael R. Shirts and John D. Chodera. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.*, 129(12):124105, 2008.

[32] 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.

[33] C. Jarzynski. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.*, 78:2690, 1997.

[34] G. Crooks. Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems. *J. Statis. Phys.*, 90:1481, 1998.

[35] 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.

[36] 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.

[37] 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.

[38] 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.