A maximum entropy thermodynamics of small systems

Purushottam D. Dixit
Biosciences Department, Brookhaven National Laboratory

We present a maximum entropy approach to analyze the state space of a small system in contact with a large bath e.g. a solvated macromolecular system. For the solute, the fluctuations around the mean values of observables are not negligible and the probability distribution $P(r)$ of the state space depends on the intricate details of the interaction of the solute with the solvent. Here, we employ a superstatistical approach: $P(r)$ is expressed as a marginal distribution summed over the variation in $\beta$, the inverse temperature of the solute. The joint distribution $P(\beta, r)$ is estimated by maximizing its entropy. We also calculate the first order system-size corrections to the canonical ensemble description of the state space. We test the development on a simple harmonic oscillator interacting with two baths with very different chemical identities viz. a) Lennard-Jones particles and b) water molecules. In both cases, our method captures the state space of the oscillator sufficiently well. Future directions and connections with traditional statistical mechanics are discussed.

* Corresponding author: Phone: (631) 344-3742; Email: pdixit@bnl.gov
I. INTRODUCTION

Recent developments in spectroscopic (1) and single molecule manipulation techniques (2, 3) allow direct measurements of very small systems. Examples of small systems from aqueous solution chemistry and biochemistry including proteins/nucleic acid molecules (4), nanoparticles (5), and hydrated ion \([\text{M(H}_2\text{O)}_n]^{k+}\) complexes interacting with bulk water solution (6). Here, we are more interested in the description of the dynamics within the solute system, e.g. folding of the protein or interaction between nanoparticles, than the details of its interactions with the surrounding medium. In other words, we seek an effective thermodynamic description of the system by integrating over the uninteresting ‘bulk’ degrees of freedom.

Due to the small size of the system and the high surface-to-volume ratio implies that the bath of solvent particles cannot be treated as an ideal thermal bath, one that interacts only weakly with the system. Consequently, the system-bath interactions have to be entertained at some detailed level of description. We expect that the behavior of a small system to be markedly different from systems with macroscopically large number of particles due to relatively heightened fluctuations in energy, volume, etc (7, 8).

To see this clearly, consider a large system coupled to a thermal bath. For simplicity, assume that the system can only exchange energy with the bath. Now imagine that the above system can be decomposed into subsystems A and B such that A is very large compared to B. For convenience, we will identify A as the solvent and B as a solute e.g. a protein. Throughout this article, we will use the terms ‘solute’ and ‘system’ and the terms ‘solvent’ and ‘bath’ interchangeably. The energy \(E\) of the composite system can be written as

\[
E(r_A, r_B) \equiv E_A(r_A) + E_B(r_B) + E_{AB}(r_A, r_B)
\]  

(1)

where \(E_A\) (\(E_B\)) is the interaction energy within the solvent (solute) and \(E_{AB}\) is the interaction between the solvent and the solute. \(r_A\) (\(r_B\)) denote the collective coordinates of solvent (solute) particles. Note that \(E \approx E_A \gg E_B(\sim E_{AB})\).

In the canonical ensemble description of the system at inverse temperature \(\beta = 1/k_B T\), we write the distribution of states \(P(r_A, r_B)\) as

\[
P(r_A, r_B | \beta) \propto \exp \left( -\beta E(r_A, r_B) \right).
\]  

(2)

The marginal distribution for the solute degrees of freedom is formally written as

\[
P(r_B | \beta) = \sum_{r_A} P(r_A, r_B | \beta) \propto \exp \left( -\beta [E_B(r_B) + \phi_B(\beta, r_B)] \right).
\]  

(3)
The marginal distribution depends only on \( r_B \), the internal coordinates of the solute, and the inverse temperature \( \beta \). Here, \( \phi_B(\beta, r_B) \) represents the temperature dependent effect of the solvent-solute interactions on the state space of the solute. Note that \( \phi_B(\beta, r_B) \) is a constant if the solute is very large, i.e. \( E_B \gg E_{AB} \), and behaves independently of the solvent. In this case, the solute acts as a thermodynamic system in its own right, in contact with an ideal thermal bath at the same inverse temperature (a thermal bath comprising of ideal gas particles).

Understanding the structure of the ‘molecular field’ \( \phi_B(\beta, r_B) \) is of tremendous importance since \( E_B(r_B) + \phi_B(\beta, r_B) \) completely describes the phase space distribution of the solute (9, 10). Even though Eq. 3 is formally true, it has little practical value since \( \phi_B(\beta, r_B) \) depends in a non-trivial fashion on the details of the solute-solvent interaction \( E_{AB} \) and, in general, is quite hard to estimate (9, 11). Ad hoc assumptions about the structure of \( \phi_B(\beta, r_B) \) are commonplace in aqueous chemistry literature including (but certainly not limited to) examples such as the Generalized Born dielectric models (12), the non-linear Poisson-Boltzmann model (13).

Here, we present a maximum entropy approach to model the state space distribution \( P(r_B) \) (see Eq. 3) by circumventing the problem of estimating the temperature dependent field \( \phi_B(\beta, r_B) \). We illustrate our framework, using molecular dynamics simulations, to study the state space of a Harmonic oscillator (B) interacting with a) a bath of Lennard-Jones particles (A) and b) a bath of water molecules (A), two solvent media of completely different chemical identities.

In sec. II we develop our method, in sec. III we explore connections with traditional statistical mechanics, and in sec. IV we illustrate it for a harmonic oscillator coupled to two realistic baths. Finally in sec. V we discuss future directions and possible connections with traditional statistical mechanics.

II. THEORY

A. Maximum entropy superstatistics

Consider a macromolecular solute interacting with a solvent as described above. The maximum entropy (maxEnt) interpretation views the problem of determining the distribution \( P(r_A, r_B) \) of the coordinates \( \{r_A, r_B\} \) of the composite system as an inference problem: \( P(r_A, r_B) \) is estimated from the limited available knowledge of the system (14). Briefly, the probabilities \( P(r_A, r_B) \) is estimated by maximizing the entropy \( S[P(r_A, r_B)] \) subject to constraints that the distribution reproduces some experimentally observed quantities such as the mean energy \( \bar{E} \). The constrained
objective function is given by Eq. 4 \((14, 15)\).

\[
S[P(r_A, r_B)] - \beta \left( \sum_{r_A, r_B} P(r_A, r_B)E(r_A, r_B) - \bar{E} \right) + \gamma \left( \sum_{r_A, r_B} P(r_A, r_B) - 1 \right).
\]

(4)

Here, \(\beta\) and \(\gamma\) are the Lagrange multipliers and

\[
S[P(r_A, r_B)] = - \sum_{r_A, r_B} P(r_A, r_B) \log P(r_A, r_B).
\]

(5)

The estimated distribution \(P(r_A, r_B|\beta)\) has the maximum entropy amongst all candidate distributions \(P^*(r_A, r_B)\) that reproduce the experimental constraints (here, average energy). \(P(r_A, r_B|\beta)\) is parameterized by a unique inverse temperature \(\beta\) and is given by Eq. 2. The maximum entropy estimate depends solely on the average energy \(\bar{E}\) and results in a correct predictive theory only for macroscopically large systems \((14, 15)\). This is due to the fact that in a macroscopically large system, the fluctuation in energy, \(\langle E^2 \rangle - \langle E \rangle^2\), is negligible compared to the mean \(\langle E \rangle\) \((14)\). Consequently, the higher moments of the energy distribution \(\langle E^n \rangle\) \((n > 1)\) can be estimated from the knowledge of the mean. A naive application of the maximum entropy principle for a small system such as the solute is bound to result in predictions that do not match with experiments.

Note if the solute \(B\) is sufficiently large, the internal interactions within the solute will vastly outweigh the solute-solvent interactions. Here, the solvent can be integrated with the thermal bath and its effect absent from the description of the solute except for setting the inverse temperature \(\beta\). In this case, the entropy of the solute \(S[P(r_B)]\) itself is maximized subject to constraining the average energy of the solute. Thus, for sufficiently large \(B\), the effect of \(A\) on the phase space of \(B\) can be represented by a single number: with respect to any prediction about the solute \(B\), the equivalence \(P(r_A, r_B|\beta) \leftrightarrow P(r_B)\) involves minimal loss of information. For a solute of intermediate size, we assume that the solvent can be equivalently represented by allowing the inverse temperature of the solute to fluctuate i.e. we work with the ansatz \(P(r_A, r_B) \leftrightarrow P(\beta, r_B)\) rather than \(P(r_A, r_B) \leftrightarrow P(r_B)\). Instead of maximizing the entropy \(S[P(r_A, r_B)]\) of \(P(r_A, r_B)\) \((16)\), we maximize the entropy of the joint distribution \(P(\beta, r_B)\). Thus, we maximize

\[
S[P(\beta, r_B)] = - \sum_{\beta, r_B} P(\beta, r_B) \log P(\beta, r_B) = S[P(\beta)] + \sum_{\beta} P(\beta)S_B(\beta)
\]

(6)

where \(S[P(\beta)]\) is the entropy of

\[
P(\beta) = \sum_{r_B} P(\beta, r_B)
\]

(7)
and

\[ S_B(\beta) = -\sum_{r_B} P(r_B|\beta) \log P(r_B|\beta) \]  \hspace{1cm} (8)

is the entropy of the solute in contact with an ideal thermal bath at inverse temperature \( \beta \) and

\[ P(r_B|\beta) \propto \exp(-\beta E_B(r_B)) \]  \hspace{1cm} (9)

is the canonical ensemble probability distribution. The ideal thermal bath can be constructed by replacing the solvent by a dilute ideal gas solvent \( A' \) at the same temperature such that \( E_B \gg E_{A'B} \).

In maximization, a straightforward choice for the constraint is \( \bar{E}_B \), the observed mean energy of the solute \( B \). Since the entropy of the joint distribution \( P(\beta, r_B) \) is maximized rather than that of the marginal distribution \( P(r_B) \), the measured entropy \( S_B \) itself becomes a constraint (17). Thus, the constrained optimization function, including the Lagrange multipliers and summing over \( r_B \) degrees of freedom, is (from Eq. 6),

\[
S[P(\beta)] + \lambda \left( \sum_{\beta} S_B(\beta)P(\beta) - \bar{S}_B \right) + \gamma \left( \sum_{\beta} P(\beta) - 1 \right) - \zeta \left( \sum_{\beta} \langle E_B(\beta) \rangle P(\beta) - \bar{E}_B \right).
\]  \hspace{1cm} (10)

Here

\[ \langle E_B(\beta) \rangle = \sum_{r_B} E_B(r_B) P(r_B|\beta) \]  \hspace{1cm} (11)

is the average energy of the solute when it is coupled to an ideal thermal bath \( (A') \) at an inverse temperature \( \beta \).

We have transformed the problem of maximizing the entropy of the joint distribution \( P(\beta, r_B) \) to the one of maximizing the entropy of \( P(\beta) \). Carrying out the maximization,

\[ P(\beta) = \frac{1}{Z(\lambda, \zeta)} \exp(\lambda S_B(\beta) - \zeta \langle E_B(\beta) \rangle) \]  \hspace{1cm} (12)

where

\[ Z(\lambda, \zeta) = \sum_{\beta} \exp(\lambda S_B(\beta) - \zeta \langle E_B(\beta) \rangle) \]  \hspace{1cm} (13)

is the generalized partition function. The marginal distribution \( P(r_B) \) is written as

\[ P(r_B) = \sum_{\beta} P(r_B|\beta) P(\beta) \]  \hspace{1cm} (14)
Eq. 12 and Eq. 14 are the main theoretical results of this work. In deriving Eq. 12 and Eq. 14, we have assumed that the effect of the bulk solvent medium \( A \) on the solute \( B \) can be captured by allowing the inverse temperature of the solute to fluctuate. Note that the above coarse graining approach implies that the marginal distribution \( P(r_B) \) belongs to a restricted family of distributions. We believe that this approach will be successful if the solute is of intermediate scale i.e. when \( \langle \delta \phi^2 \rangle \) is small but cannot be completely neglected.

B. Implicit solvation

In the development presented here, notice that the predictions about the solute state space (Eq. 12 and Eq. 14) are independent of the details of the solute-solvent interactions. From the perspective of effective interaction models, the current development closely resembles an implicit solvation model.

The framework assumes that solvent induced modulation in the state space of the solute is completely characterized by allowing the temperature of the solute to vary. The distribution \( P(\beta) \) of the temperature \( \beta \) of the solute is governed by \( \lambda \) and \( \zeta \). In short, \( \lambda \) and \( \zeta \) (along with \( S_B(\beta) \) and \( \langle E_B \rangle_\beta \); both are properties of the solute only) completely describes solute-solvent interactions. This is the result of the coarse graining where \( P(r_A, r_B) \) is assumed to be equivalent to \( P(\beta, r_B) \) with respect to all predictions about the solute \( B \). We believe that Eq. 12 and Eq. 14 present an implicit solvation model that is independent of the details of the solute-solvent chemistry and has a rigorous basis in the maximum entropy framework. We leave it for further studies to study more realistic systems using the current framework.

III. CONNECTIONS TO STATISTICAL MECHANICS

It is instructive to examine the limiting behavior of Eq. 12. Let us first write \( \zeta = \beta_0 \lambda \). The maximum of the \( P(\beta) \) distribution can be found out by setting the first \( \beta \) derivative of \( \lambda S_B(\beta) - \beta_0 \lambda \langle E_B \rangle_\beta \) to zero. Differentiating with respect to \( \beta \), we get

\[
-\lambda \frac{C_v(\beta)}{\beta} + \lambda \beta_0 \frac{C_v(\beta)}{\beta^2} = 0
\]

It’s easy to check that the second derivative is positive at \( \beta = \beta_0 \). Thus, if the heat capacity \( C_v(\beta) \) and \( \lambda \) are non-zero, the maximum of the \( P(\beta) \) distribution lies at \( \beta = \beta_0 \).

For large \( \lambda \) keeping \( \beta_0 \) finite, \( P(\beta) \) will tend to a narrowly peaked distribution around \( \beta = \beta_0 \) and can be described by a normal distribution. The width of the normal distribution is dictated
by the second derivative. If \( C_v(\beta) \) doesn’t vary rapidly around \( \beta = \beta_0 \), we can show that for large \( \lambda \),

\[
P(\beta) \sim \mathcal{N} \left( \beta_0, \frac{\sqrt{2} \beta_0}{\lambda C_v(\beta_0)} \right).
\]  

(15)

Thus, as \( \lambda \to \infty \), the state space of the solute \( B \) is described by a canonical ensemble distribution

at inverse temperature \( \beta = \beta_0 = \zeta/\lambda \). Note that as \( \lambda \to \infty \), \( \beta_0 \) is also the temperature of the surrounding bath. We conclude that \( \lambda < \infty \) captures the strength of the solute-solvent interactions and \( \beta_0 = \zeta/\lambda \) is the effective temperature of the solute.

For weak solute-solvent interactions, i.e. large \( \lambda \), the first order corrections in solute behavior due to solute-solvent interactions, i.e. the first order estimate of \( \phi(\beta_0, r_B) \) (see Eq. 3), can be computed if one knows the heat capacity of the solute at \( \beta_0 \), the effective temperature of the solute. Dropping the \( \beta_0 \) dependence of \( C_v \) for brevity and writing \( P(r_B) \)

\[
P(r_B) = \int P(\beta) P(r_B|\beta) d\beta \approx \int_0^{\infty} \mathcal{N} \left( \beta_0, \frac{\sqrt{2} \beta_0}{\lambda C_v(\beta_0)} \right) \exp \left( -\beta E(r_B) + \beta F(\beta_0) \right) d\beta
\]

\[
\propto \exp \left( -\beta_0 E(r_B) \right) \exp \left( \frac{\beta_0^2 E(r_B)^2}{\lambda C_v} \right) \left( 1 + \text{Erf} \left[ \frac{\lambda C_v - 2 \beta_0 E(r_B)}{2 \sqrt{\pi C_v}} \right] \right)
\]

(16)

(17)

From Eq. 17 by expanding for energies small compared to \( \lambda C_v \), we estimate \( \phi(\beta_0, r_B) \) as (see Eq. 3 for the definition of \( \phi \))

\[
\phi(\beta_0, r_B) \approx e^{-\frac{C_v \lambda}{\pi} E(r_B)} - \frac{\beta_0 E(r_B)^2}{C_v \lambda}
\]

(18)

As expected, \( \phi(\beta_0, r_B) \to 0 \) when \( \lambda \to \infty \). It will be interesting to explore if Eq. 17 and Eq. 18 can be exploited in molecular dynamics simulations. We leave it for further studies.

Even under the simplifying condition of weak solute-solvent interactions, i.e. large \( \lambda \), the maximum entropy approach suggests that the solvent modulates the phase space behavior of the solute in a highly non-trivial manner. Eq. 17 and Eq. 18 provide a glimpse into the mechanism of the solvent-induced modulation of the phase space of the solute. The first term effectively decreases its temperature while the second term increases the propensity to sample high energy states thus effectively increasing the temperature.

Notice that the phase space described by Eq. 18 is identical to a maxEnt probability distribution where the mean energy \( \langle E \rangle \) and the fluctuation in the energy \( \langle E^2 \rangle \) are constrained (see Eq. 18 and
Eq. 3. Previously, the predictions from Eq. 18 have been validated for ferromagnetic materials (18, 19). In the current framework however, instead of constraining the higher moments of energy, we have constrained the average entropy of the small system. The average entropy is a natural constraint since due to non-negligible system-bath interactions unlike for a thermodynamic system, the entropy of the distribution \( P(r_B) \) of the phase space of the small system will not be at its maximum under the constraint of mean energy. We speculate that if the ansatz \( P(r_A, r_B) \leftrightarrow P(\beta, r_B) \) (see above) is violated, constraining the higher moments of the energy may prove useful.

IV. HARMONIC OSCILLATOR

A. Theory

We analytically illustrate the above development (Eq. 12, Eq. 14, and Eq. 18), using molecular dynamics simulations (20), for a harmonic oscillator, comprising of two Lennard-Jones particles, coupled to a solvent bath of a) hydrophobic Lennard-Jones particles and b) water molecules.

Imagine a harmonic oscillator solute (B) interacting with a bath of solvent particles (A) at inverse temperature \( \beta \). Without loss of generality, let the energy of the oscillator be \( E(r) = r^2 \). If the interactions of the oscillator with the surrounding particles are very weak, we know that the probability distribution \( P(r) \) is parametrized by the inverse temperature and is given by

\[
P(r|\beta) = \frac{4r^2\beta^{3/2}e^{-r^2\beta}}{\sqrt{\pi}}. \tag{19}
\]

Note that for the harmonic oscillator in contact with a weakly interacting bath, the average energy \( \langle E \rangle_\beta \sim 1/\beta \) and the entropy \( S(\beta) \sim \log \beta \) upto an additive and multiplicative constant.

If the oscillator-solvent interactions are not weak, Eq. 19 does not adequately describe \( P(r) \) (see Fig. 2). In this case, as discussed above, we allow the inverse temperature \( \beta \) of the oscillator to fluctuate. From Eq. 12

\[
P(\beta) \propto \exp \left( \lambda \log \beta - \frac{\zeta}{\beta} \right)
\]

\[
= e^{-\frac{\zeta}{\beta} - \lambda \zeta^{\lambda-1}} \frac{1}{\Gamma(\lambda-1)}. \tag{20}
\]

Before we derive the marginal distribution \( P(r) \), let us inspect the limiting behavior of \( P(\beta) \). Figure 1 shows \( P(\beta) \) for different values of \( \lambda \) and \( \zeta \), keeping \( \beta_0 = \zeta/\lambda \) fixed. Observe that at fixed \( \beta_0 \), \( P(\beta) \) distributions with higher values of \( \lambda \) is peaked and becomes a dirac delta distribution \( \delta(\beta - \zeta/\lambda) \). Thus, as seen above, while \( \beta_0 = \zeta/\lambda \) represents the temperature of the solute, \( \lambda \)
FIG. 1. The distribution $P(\beta)$ of the temperature of the harmonic oscillator solute. As $\lambda \to \infty$, $P(\beta)$, the distribution of $\beta$ approaches a Dirac delta distribution. Notice that at small values of $\lambda$, the $P(\beta)$ distribution is very broad. Hence, while $\zeta/\lambda$ is equivalent to the temperature of the solute, $\lambda$ captures the strength of the interactions between the solvent and the solute.

characterizes the strength of its interactions with the solvent: higher $\lambda$ implies weaker interactions. We hypothesize that in an experiment where the solvent composition is fixed, $\lambda$ cannot be changed smoothly while $\beta_0 = \zeta/\lambda$ can be tuned by changing the temperature of the bath; however, in case of an aqueous solvent, $\lambda$ can be changed by perturbing the solute-solvent interactions using osmolytes and salts.

Finally, the marginal distribution $P(r) = \int P(\beta) \cdot P(r|\beta) d\beta$ is given by,

$$P(r) = \frac{8r^{\lambda-\frac{1}{2}}\zeta^{\frac{1}{2}+\frac{1}{4}}K_{\lambda-\frac{5}{2}}(2r\sqrt{\zeta})}{\sqrt{\pi}\Gamma(\lambda-1)}$$  \hspace{1cm} (21)

Here, $K_{\gamma}(x)$ is the modified Bessel function of the second kind with parameter $\gamma$. To understand Eq. 21 physically, let’s write $\zeta = \beta_0 \lambda$ and calculate the average energy

$$\langle E \rangle = \langle r^2 \rangle = 3\frac{(\lambda-1)}{2\beta_0 \lambda}$$  \hspace{1cm} (22)

As $\lambda \to \infty$,

$$\langle E \rangle = \frac{3}{2\beta_0}$$  \hspace{1cm} (23)

Hence, as $\lambda \to \infty$, the oscillator behaves as if it is in contact with an ideal thermal bath at inverse temperature $2\beta_0/3$. Similarly, it is easy to check that the limiting behavior of Eq. 21 is equal to that of Eq. 12, the distribution of states of a harmonic oscillator in contact with an ideal thermal bath.
FIG. 2. The experimentally observed \( P(r) \) distribution (black circles) compared with the best-fit canonical ensemble distribution (black line, Eq. [19] and the best-fit of Eq. [21] (red line). Left panel shows the harmonic oscillator interacting with water molecules and the right panel shows the oscillator interacting with Lennard-Jones particles. Note that these two solvent systems have very different chemical identities. Yet, Eq. [21] describes the experimentally observed distribution over 5 orders of magnitude especially in the extended tail that the canonical ensemble distribution Eq. [19] fails to capture.

B. Numerical simulation

An interesting aspect of the development above is that it hypothesizes that the form of the probability distribution \( P(r_B) \) independent of the chemistry of the solute-solvent interactions. The solvent medium affects the solute only through the ‘coupling parameter’ \( \lambda \) (see Eq. [12] and Eq. [14]). We test this prediction by studying a harmonic oscillator interacting with two systems that are chemically very different from each other viz. a bath of Lennard-Jones particles and a bath of water molecules (see appendix for details).

Fig. 2 shows the observed probability distribution (black circles) from the molecular dynamics simulations, the best-fit canonical ensemble distribution (black line, see Eq. [19]), and the best-fit of Eq. [21] (see appendix for the fitting procedure). Note that even though these two solvent systems have very different chemical identities, fig. 2 clearly shows that the empirically observed distribution for the two systems has an extended tail that cannot be captured by the canonical ensemble distribution. Meanwhile, the distribution \( P(r) \) of Eq. [21] captures the empirically observed distribution extremely well.
V. DISCUSSION

In traditional statistical mechanics, the inverse temperature $\beta$ is an intensive variable and its fluctuation does not have a direct physical interpretation. Here, the entropy is related to heat loss and is a measurable quantity. On the other hand, the maximum entropy (maxEnt) interpretation of statistical mechanics finds its basis in the information theoretic interpretation of entropy \cite{14,15,21} where the inverse temperature $\beta$ is a Lagrange multiplier and entropy is an inference tool. Within the maxEnt framework $\beta$ can be treated as a parameter for the canonical ensemble distribution $P(r_B|\beta)$. Fluctuations in $\beta$ can be interpreted as arising due to the hyper-probability $P(\beta)$ in the parameter space \cite{22}.

The maxEnt interpretation views statistical mechanics as an inference problem: amongst candidate distributions $P^*(r_B)$ that reproduce the known experimental measurements about the system $B$, maxEnt states that the one that has the maximum entropy describes the internal states of the system correctly. For a system exchanging energy with the bath, the success of maxEnt is heavily dependent on the assumption of weak system-bath interactions (i.e. weak solute-solvent interactions) \cite{14,15}. Here, we have shown that the maxEnt interpretation also offers a way to generalize statistical mechanics.

Note that regardless of the size of the system $B$, the entropy of the joint system comprising of $B$ and the bath $A$ is maximized \cite{16}. When the interactions between $A$ and $B$ are not too strong (but not too weak to be negligible either), we hypothesized that the effect of the variation in the states $r_A$ of the solvent bath on the solute can be characterized by a varying its inverse temperature $\beta$. In other words, we assumed that the equivalence $P(r_A, r_B) \leftrightarrow P(\beta, r_B)$ involved minimal loss of information about the bath. Our development lead to a generalization of statistical mechanics that resembles superstatistics \cite{17,23}. We established a direct correspondence between the developed framework and traditional statistical mechanics in the limiting case of very weak solute-solvent interactions. We also calculated the first order correction to canonical ensemble description of the state space of small systems.

We illustrated the framework with a harmonic oscillator coupled to two solvent baths of very different chemical identities. We showed that the superstatistical distribution Eq. \ref{eq:superstatistical} describes the distribution of states $P(r)$ of the oscillator better than the usual canonical ensemble distribution Eq.\ref{eq:canonical}. The current development has shown that the thermodynamics of small systems can be suitably described by a superstatistics. We believe that the framework will be useful in understanding solvent induced modulations of solute state space e.g. with implicit solvent models.
VI. ACKNOWLEDGMENT

I would like to thank Prof. Ken Dill, Prof. Dilip Asthagiri, Prof. German Drazer, and Mr. Sumedh Risbud for stimulating conversations and suggestions about the manuscript.

This work was supported by grants PM-031 from the Office of Biological Research of the U.S. Department of Energy.

[1] Roy, R., S. Hohng, and T. Ha, 2008. A practical guide to single-molecule FRET. Nature methods 5:507–516.
[2] Greenleaf, W., M. Woodside, and S. Block, 2007. High-resolution, single-molecule measurements of biomolecular motion. Annual review of biophysics and biomolecular structure 36:171.
[3] Mehta, A., M. Rief, J. Spudich, D. Smith, and R. Simmons, 1999. Single-molecule biomechanics with optical methods. Science 283:1689–1695.
[4] Garrett, R., and C. Grisham, 1999. Biochemistry. Saunders College Publishing.
[5] Rao, C., A. Müller, and A. Cheetham, 2006. The chemistry of nanomaterials, volume 1. Wiley-VCH.
[6] Marcus, Y., 1994. A simple empirical model describing the thermodynamics of hydration of ions of widely varying charges, sizes, and shapes. Biophysical chemistry 51:111–127.
[7] Hill, T. L., 2002. Thermodynamics of small systems. Dover publications, Mineola, New York.
[8] Hill, T. L., 2001. A different approach to nanothermodynamics. Nano Lett. 1:273–275.
[9] Dixit, P. D., and D. Asthagiri, 2011. The role of bulk protein in local models of ion-binding to proteins. Comparative study of KcsA, its semi-synthetic analog with a locked-in binding site, and Valinomycin. Biophys. J. 100:1542–1549.
[10] Dixit, P., and D. Asthagiri, 2011. An elastic-network-based local molecular field analysis of zinc finger proteins. J. Phys. Chem. B 115:7374–7382.
[11] Dixit, P. D., and D. Asthagiri, 2011. Thermodynamics of ion selectivity in the KcsA K$^+$ channel. The Journal of general physiology 137:427–433.
[12] Bashford, D., and D. Case, 2000. Generalized Born models of macromolecular solvation effects. Ann. Rev. Phys. Chem. 51:129–152.
[13] Sharp, K., and B. Honig, 1990. Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. J. Phys. Chem 94:7684–7692.
[14] Jaynes, E. T., 1957. Information theory and statistical mechanics I. Phys. Rev. 106:620–630.
[15] Pressé, S., K. Ghosh, J. Lee, and K. A. Dill, 2012. The principles of Maximum Entropy and Maximum Caliber in statistical physics. Rev. Mod. Phys. to appear.
[16] Lee, J., and S. Pressé, 2012. Microcanonical origin of the maximum entropy principle for open systems. Phys. Rev. E to appear.
[17] Crooks, G. E., 2008. Beyond Boltzmann-Gibbs statistics: Maximum entropy hyperensembles out-of-equilibrium. *Phys. Rev. E* 75:041119.

[18] Chamberlin, R. V., J. V. Vermaas, and G. H. Wolf, 2009. Beyond the Boltzmann factor for corrections to scaling in ferromagnetic materials and critical fluids. *The European Physical Journal B* 71:1–6.

[19] Chamberlin, R. V., and G. H. Wolf, 2009. Fluctuation-theory constraint for extensive entropy in Monte-Carlo simulations. *The European Physical Journal B* 67:495–499.

[20] Phillips, J., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. Skeel, L. Kale, and K. Schulten, 2005. Scalable Molecular Dynamics with NAMD. *J. Comp. Chem.* 26:1781–1802.

[21] Shore, J., and R. Johnson, 1980. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Info. Theory* 26:26–37.

[22] Caticha, A., and R. Preuss, 2004. Maximum entropy and Bayesian data analysis: Entropic prior distributions. *Phys. Rev. E* 70:046127.

[23] Beck, C., and E. G. D. Cohen, 2003. Superstatistics. *Physica A* 322:267–275.

[24] Jorgensen, W., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys* 79:926–935.

[25] Neria, E., S. Fischer, and M. Karplus, 1996. Simulation of activation free energies in molecular systems. *J. Chem. Phys* 105:1902–1921.

[26] MacKerell, Jr., A. D., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, and D. Joseph-McCarthy, 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* 102:3586–3616.
A harmonic spring consisting of two Lennard-Jones particles was immersed in a bath of 512 Lennard-Jones particles in a cube of side 25Å and a bath of 256 TIP3 (24, 25) water molecules at 300. NVT molecular dynamics simulations were run with NAMD (20) at 300K. The CHARMM (26) forcefield was used to describe the interaction between the Lennard-Jones particles and between the spring and the bath of particles. The spring constant for the harmonic oscillator was chosen to be \( k = 0.5 \text{ kcal/mol} \cdot \text{Å}^2 \). The \( \epsilon \) parameter for the Lennard-Jones bath was set at \( -0.015 \) while the \( \epsilon \) parameter for the spring was set at \( \epsilon = -10.0 \). The size parameter was set at \( r = 2.1\text{Å} \) for the oscillator particles (\( r = 1.1\text{Å} \) when interacting with water) and \( r = 1.1\text{Å} \) for the bath particles. The systems were minimized for 2000 steps followed by an equilibration of 1 nanosecond and a production run of 2 nanosecond. Configurations were stored every 100 femtosecond.

B. Fitting the parameters

In order to fit Eq. 21 to the experimental data, one needs to determine the free parameters from the simulation. In the traditional canonical ensemble, the inverse temperature \( \beta \) of the harmonic oscillator will be estimated from its average energy. Here, we show how to estimate the free parameters from the simulation. It is non-trivial to measure the average system entropy \( \langle S_B(\beta) \rangle \) in a computer simulation. Yet, operationally,

\[
\langle S_B(\beta) \rangle \propto \int \log \beta \cdot P(\beta) d\beta
\]

\[
\propto \int \int \log r \cdot P(r|\beta) P(\beta) dr d\beta
\]

\[
= \int \langle \log r \rangle_{\beta} P(\beta) d\beta.
\]

In other words, constraining \( S_B(\beta) \) is equivalent to constraining \( \log r \). Thus, in addition to \( \langle E \rangle = \langle r^2 \rangle \), we also estimate \( \langle \log r \rangle \) from the simulation and then fit Eq. 21.