

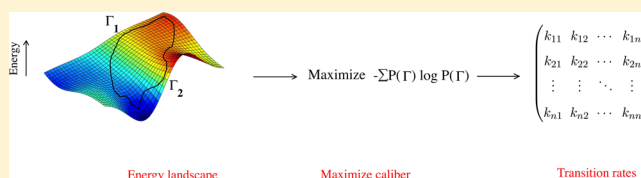
Inferring Microscopic Kinetic Rates from Stationary State Distributions

Purushottam D. Dixit^{*,†} and Ken A. Dill[†]

[†]Department of Systems Biology, Columbia University, New York, New York 10032, United States

[‡]Laufer Center for Physical and Quantitative Biology, Departments of Chemistry, and Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794, United States

ABSTRACT: We present a principled approach for estimating the matrix of microscopic transition probabilities among states of a Markov process, given only its stationary state population distribution and a single average global kinetic observable. We adapt Maximum Caliber, a variational principle in which the path entropy is maximized over the distribution of all possible trajectories, subject to basic kinetic constraints and some average dynamical observables. We illustrate the method by computing the solvation dynamics of water molecules from molecular dynamics trajectories.



INTRODUCTION

We propose a method for inferring the kinetic rate matrix for stochastic systems for which the steady state populations are known. The types of systems of interest include protein folding,^{1–3} ion channels,⁴ molecular motors,⁵ the evolutionary dynamics of protein sequences,⁶ the collective firing patterns of neurons,⁷ or noisy gene expression.⁸ In such studies, stationary probabilities $\{p_i\}$ of N stable states $\{i\}$ are often known or estimated from experimental data,^{6,7} detailed molecular mechanics calculations,³ coarse grained theories,^{1,4} or from Maximum Entropy.^{6,7,9–12} Moreover, suppose further that we know some average global dynamical quantity, such as the average mutation rate (number of amino acid changes per unit time), the average current through an ion channel, or the average number of neurons that change their state from spiking to resting (and vice versa) per unit time step. How can we infer the full distribution of microscopic state-to-state rates, given just the stationary-state populations and one or more average overall rate quantities?

More specifically, we are interested in a principled way to solve the following under-determined “inverse” kinetics problem. Consider a stationary and irreducible Markov process among $i = 1, 2, 3, \dots, N$ states. Suppose you know the following: (a) the stationary state probability distribution, $\{p_i\}$ of the occupancies of those states, and (b) the value $\langle w \rangle$ of some dynamical observable w averaged over the ensemble of stationary state trajectories. From these $N + 1$ quantities, we want to infer the $N \times N$ microscopic transition probabilities,

$$k_{ij} = P(X(t + dt) = j | X(t) = i) \quad (1)$$

between those states.

While the stationary distribution indeed constrains the transition probability matrix, it does not uniquely determine it. Here, we develop a procedure based on the principle of Maximum Caliber, a variant of the principle of Maximum

Entropy, that is applicable to dynamical processes.^{13,14} This variational principle allows us to uniquely infer $\sim N^2$ transition probabilities that are consistent with $\sim N$ imposed constraints.

First, we define *the path entropy*, S , over a given ensemble $\{\Gamma\}$ of trajectories $\Gamma \equiv \dots \rightarrow i \rightarrow j \rightarrow k \rightarrow l \dots$ as

$$S = - \sum_{\{\Gamma\}} p(\Gamma) \log p(\Gamma) \quad (2)$$

Maximum Caliber is a variational principle that chooses a unique probability distribution $\{P(\Gamma)\}$ over trajectories from all possible candidate distributions as the one that maximizes the path entropy while otherwise satisfying the relevant stationary and dynamical constraints.^{13–16}

Consider an ensemble of stationary state trajectories $\{\Gamma\}$ (see above) having a total time duration T . Restricting our attention to first-order Markov trajectories allows us to greatly simplify the path entropy and carry out our analysis in terms of trajectories of single steps, $i \rightarrow j$.^{2,17–19} The Markov property implies that the probability of any particular trajectory Γ can be expressed in terms of the transition probabilities $\{k_{ij}\}$,

$$P(\Gamma) = \dots k_{ij} \cdot k_{jk} \cdot k_{kl} \dots \quad (3)$$

The path entropy of the above ensemble is directly proportional to the total duration T of the trajectory. The path entropy per unit time is given by^{20,21}

$$S = - \sum_{ij} p_i k_{ij} \log k_{ij} \quad (4)$$

The microscopic transition probabilities of any Markov process are subject to two types of constraints. First, from state i at time t , the system *must land somewhere* at time $t + dt$.

Received: February 17, 2014

Published: June 2, 2014

Second, a system in state j at time $t + dt$ must arrive from somewhere, so

$$\sum_j k_{ij} = 1 \quad \forall i \quad \text{and} \quad p_j = \sum_i p_i k_{ij} \quad \forall j \quad (5)$$

Third, we require one additional constraint that is global, that is, averaged over the entire ensemble of trajectories. We fix the path ensemble average of some dynamical quantity w . The average $\langle w \rangle_\Gamma$ over any given stationary trajectory Γ is given by

$$\langle w \rangle_\Gamma = \frac{1}{T} (\dots + w_{ij} + w_{jk} + w_{kl} + \dots) \quad (6)$$

The path ensemble average $\langle w \rangle$ is

$$\langle w \rangle = \sum_\Gamma P(\Gamma) \langle w \rangle_\Gamma \quad (7)$$

Since Γ is a stationary state trajectory, the path ensemble average $\langle w \rangle$ of eq 7 simplifies to

$$\langle w \rangle = \sum_{i,j} p_i k_{ij} w_{ij} \quad (8)$$

Maximization of the path entropy subject to these three constraints can be expressed equivalently in terms of maximization of a quantity called the Caliber C :¹⁴

$$C = - \sum_{i,j} p_i k_{ij} \log k_{ij} + \sum_i a_i (\sum_j p_i k_{ij} - p_i) + \sum_j l_j (\sum_i p_i k_{ij} - p_j) - \gamma (\sum_{i,j} p_i k_{ij} w_{ij} - \langle w \rangle) \quad (9)$$

where γ is the Lagrange multiplier associated with the constraint $\langle w \rangle$ and $\{a_i\}$ and $\{l_j\}$ enforce the *to-somewhere* constraint and the *from-somewhere* constraint, respectively.

To solve for the matrix k_{ij} of transition probabilities, we take the derivative of the Caliber C with respect to k_{ij} and equate it to zero. This gives

$$p_i (1 + \log k_{ij}^*) = a_i p_i + l_j p_j - \gamma p_i w_{ij} \Rightarrow k_{ij}^* = \frac{\beta_i}{p_i} \lambda_j e^{-\gamma w_{ij}} \quad (10)$$

where we have made the substitutions: $e^{a_i-1} = \beta_i/p_i$ and $e^{l_j} = \lambda_j$. The values k_{ij}^* are the transition probabilities that satisfy the constraints and otherwise maximize the caliber. For simplicity of notation, we drop the superscript $*$ in the remainder of this paper, that is, $k_{ij}^* \equiv k_{ij}$.

In this problem, the values of p_i are given. To compute the k_{ij} values, we first must determine the values of the Lagrange multipliers β_i , λ_j , and γ . We do so by substituting the constraint relations mentioned above.

Determining the Lagrange Multipliers. For a given value of γ , the modified Lagrange multipliers β_i and λ_j are determined by satisfying the *to-somewhere* and *from-somewhere* conditions indicated above. From eqs 5

$$\begin{aligned} \sum_j \frac{\beta_i}{p_i} \lambda_j W_{ij} &= 1 \quad \text{and} \\ p_j &= \sum_i p_i \frac{\beta_i}{p_i} \lambda_j W_{ij} \Rightarrow \frac{p_j}{\beta_i} = \sum_j \lambda_j W_{ij} \quad \text{and} \\ \frac{p_j}{\lambda_j} &= \sum_i \beta_i W_{ij} \end{aligned} \quad (11)$$

where $e^{-\gamma w_{ij}} = W_{ij}$. Equation 11 can be simplified if we define a nonlinear operator D over column vectors $\bar{x} = [x_1, x_2, \dots]^T$ as $D[\bar{x}]_i = p_i/x_i$. We have

$$W\bar{\lambda} = D[\bar{\beta}] \quad \text{and} \quad W^T\bar{\beta} = D[\bar{\lambda}] \quad (12)$$

where $\bar{\lambda} = [\lambda_1, \lambda_2, \dots]^T$ and $\bar{\beta} = [\beta_1, \beta_2, \dots]^T$ are the column vectors of Lagrange multipliers.

For a particular value of the Lagrange multiplier γ , eqs 12 can be numerically and self-consistently solved for $\{\beta_i\}$ and $\{\lambda_j\}$. In practice, we choose an appropriate γ by first constructing transition probabilities $\{k_{ij}\}$ for multiple values of γ (see eq 10) and choosing the value of γ that satisfies

$$\langle w \rangle = \sum_{i,j} p_i k_{ij} w_{ij} = \sum_{i,j} \beta_i \lambda_j e^{-\gamma w_{ij}} w_{ij} \quad (13)$$

where $\langle w \rangle$ is the prescribed value of the ensemble average of the dynamical quantity w .

Detailed Balance. So far, the treatment above is generally applicable to nonequilibrium systems. However, if we are interested in restricting our attention to systems in thermodynamic equilibrium, we can impose an additional constraint that the system must satisfy detailed balance, $p_i k_{ij} = p_j k_{ji}$. In this case, the Caliber can be expressed as

$$\begin{aligned} C &= - \sum_{i,j} p_i k_{ij} \log k_{ij} + \sum_i a_i (\sum_j p_i k_{ij} - p_i) \\ &+ \sum_j l_j (\sum_i p_i k_{ij} - p_j) - \gamma (\sum_{i,j} p_i k_{ij} w_{ij} - \langle w \rangle) \\ &+ \sum_{i,j} d_{ij} (p_i k_{ij} - p_j k_{ji}) \end{aligned} \quad (14)$$

Here, d_{ij} are Lagrange multipliers that impose the detailed balance condition. Differentiating with respect to k_{ij} and setting the derivative to zero, we find (see eq 10)

$$\begin{aligned} p_i (1 + \log k_{ij}) &= a_i p_i + l_j p_j - \gamma p_i w_{ij} + p_i (d_{ij} - d_{ji}) \Rightarrow k_{ij} \\ &= \frac{\beta_i}{p_i} \lambda_j e^{-\gamma w_{ij}} \delta_{ij} \end{aligned} \quad (15)$$

where $\delta_{ij} = e^{d_{ij}-d_{ji}} = 1/\delta_{ji}$.

Now we have to determine the modified Lagrange multipliers β_i , λ_j , δ_{ij} , and γ from the imposed constraints. Let us first impose detailed balance to determine δ_{ij} . We have

$$p_i k_{ij} = p_i \frac{\beta_i}{p_i} \lambda_j e^{-\gamma w_{ij}} \delta_{ij} = \beta_i \lambda_j e^{-\gamma w_{ij}} \delta_{ij} = p_j k_{ji} = \beta_j \lambda_i e^{-\gamma w_{ji}} \delta_{ji} \quad (16)$$

Since $\delta_{ij} = 1/\delta_{ji}$, we have

$$\delta_{ij} = \sqrt{\frac{\beta_j \lambda_i}{\beta_i \lambda_j}} e^{+\gamma w_{ij} - \gamma w_{ji}} \Rightarrow k_{ij} = \frac{1}{p_i} \sqrt{\beta_i \lambda_i \beta_j \lambda_j} e^{-\gamma/2(w_{ij} + w_{ji})} \quad (17)$$

In eq 17, identifying $\rho_i = (\beta_i \lambda_i)^{1/2}$,

$$k_{ij} = \frac{\rho_i \rho_j}{p_i} W_{ij}^{\text{sym}} \quad (18)$$

where $e^{-\gamma/2(w_{ij} + w_{ji})} = W_{ij}^{\text{sym}}$ is the symmetrized form of $W_{ij} = e^{-\gamma w_{ij}}$ (see above). It is easy to see that k_{ij} in eq 18 satisfies detailed balance. The Lagrange multipliers $\bar{\rho} = [\rho_1, \rho_2, \dots, \rho_N]$ and γ are determined by the same procedure as described above. We have

$$\sum_j k_{ij} = 1 \Rightarrow \sum_j \frac{\rho_i}{\rho_j} W_{ij}^{\text{sym}} = 1 \Rightarrow \sum_j W_{ij}^{\text{sym}} \rho_j = \frac{\rho_i}{\rho_i} \quad (19)$$

In other words, $\bar{\rho}$ is the solution of the nonlinear equation

$$W^{\text{sym}} \bar{\rho} = D[\bar{\rho}] \quad (20)$$

and γ is determined by adjusting the path ensemble average

$$\langle w \rangle = \sum_{i,j} p_i k_{ij} w_{ij} = \sum_{i,j} \rho_i \rho_j e^{-\gamma/2(w_{ij}+w_{ji})} w_{ij} \quad (21)$$

■ AN ILLUSTRATION: COMPUTING THE DYNAMICS OF A SOLVATION SHELL FROM SIMULATED POPULATIONS

We now illustrate how the present MaxCal method can be used to take a stationary-state distribution and a global dynamical constraint and to infer microscopic kinetics. Consider a shell of solvating water molecules surrounding a single water molecule. The number, $n(t)$, of water molecules in the hydration shell is a quantity that fluctuates with time t (see Figure 1). We

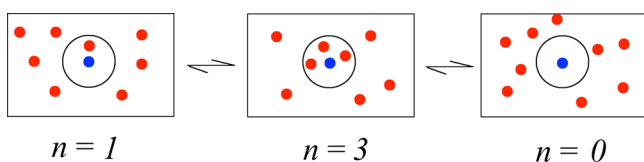


Figure 1. Hydration shell (black circle) around a central water molecule (blue disc) is dynamically populated by other water molecules in the bulk solvent medium (red discs). The probability, $p(n)$ that the hydration shell has exactly n water molecules is a key quantity in determining the solvation free energy of liquid water.^{23,24}

want to compute how fast the water molecules enter or exit the solvation shell. If the time interval dt is small, $n(t)$ and $n(t+dt)$ will be statistically correlated. Here, we construct a Markov process to model the time series $\{n(t)\}$. We will require the Markov process to reproduce (a) the stationary distribution $p(n)$ that is observed in molecular dynamics simulations, and (b) the average change in occupancy Δ per time step of duration dt , a path ensemble average (see Appendix for details of molecular simulation). While the choice of dynamical constraint(s) remains somewhat arbitrary, it is validated only *a posteriori*. In other words, the constrain(s) is a modeling aspect of any maximum entropy method.²² We have

$$\Delta = \langle |n(t+dt) - n(t)| \rangle = \sum_{i,j} |i-j| p(i) k_{ij} \quad (22)$$

where $n(t) = i$ and $n(t+dt) = j$.

Since the system should satisfy detailed balance, the transition probability k_{ij} for a transition $n(t) = i \rightarrow n(t+dt) = j$ is given by (see eq 18),

$$k_{ij} = \frac{\rho_i \rho_j e^{-\gamma|i-j|}}{p(i)} \quad (23)$$

For a given value of γ , we determine the Lagrange multipliers ρ_i from eqs 20 above. In order to determine the Lagrange multiplier γ which dictates the rate of transition between states, we first construct Markov processes for different values of γ . Panel B of Figure 2 shows that the path ensemble average of the change in occupation number per unit time step

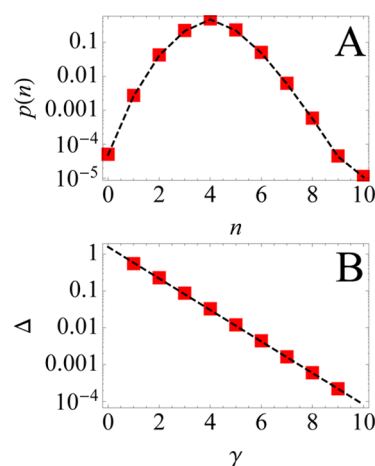


Figure 2. Panel A: The stationary distribution $p(n)$ of the number of water molecules in the hydration shell of radius $r = 3.2$ Å of a reference water molecule. Panel B: The dependence of the ensemble average of change in water occupancy number $\Delta = \langle |n(t+dt) - n(t)| \rangle$ on the Lagrange multiplier γ . We see that Δ depends exponentially on γ . A higher γ implies slower dynamics and vice versa. We choose $\gamma = 3.29$ to match the observed $\Delta \approx 0.0629$ in the molecular dynamics simulation.

Δ is exponentially decreasing with γ . From trajectories sampled at every 5 fs from the MD simulation, we find that experimental trajectory average $\Delta_{\text{expt}} = \overline{|n(t+dt) - n(t)|} \approx 0.0629$ which corresponds to $\gamma \approx 3.29$. From here onward, we use $\gamma = 3.29$ and construct the transition probabilities $\{k_{ij}\}$ (see eq 23). Note that the path ensemble average Δ and consequently the Lagrange multiplier γ , depend on the time interval dt between two observation ($dt = 5$ fs here).

From the Markov process constructed with $\gamma = 3.29$ (see above), we now compute various dynamical quantities: (a) the probability P_d of jump size d , (b) the occupancy autocorrelation $\langle \delta_{n(0),n(\tau)} \rangle$, and (c) the transition probabilities k_{ij} , and we compare to those obtained directly from the MD simulation trajectory. In general, the MaxCal method will be of value when transitions are hard to simulate, such as for large kinetic barriers. Here, we are just illustrating with a toy problem for which we can determine the transition probabilities independently from the simulations.

From the long simulation trajectory, the probability P_d of jump size is estimated as the histogram of $d = n(t+dt) - n(t)$. Here d could be both positive and negative. P_d is given by

$$P_d = \sum_n p(n) k_{n,n+d} \quad (24)$$

The normalized occupancy autocorrelation is simply the joint probability that $n(t)$ and $n(t+\tau)$ are equal. It is given by

$$\langle \delta_{n(0),n(\tau)} \rangle = \frac{\sum_n p(n) K(\tau)_{nn}}{\sum_n p(n)^2} \quad (25)$$

where $K(\tau) = k^\tau$ is the τ^{th} power of the matrix of transition probabilities $\{k_{ij}\}$.

In Figure 3 we plot P_d , $\langle \delta_{n(0),n(\tau)} \rangle$, and $\{k_{ij}\}$ estimated from the molecular dynamics trajectory and compare them to our predictions from Markov modeling. Even though we constrained only the mean value $\Delta = \langle |d| \rangle$ of P_d , the modeled Markov process captures the entire distribution P_d with high accuracy. Similarly the occupancy correlation $\langle \delta_{n(0),n(\tau)} \rangle$ is also

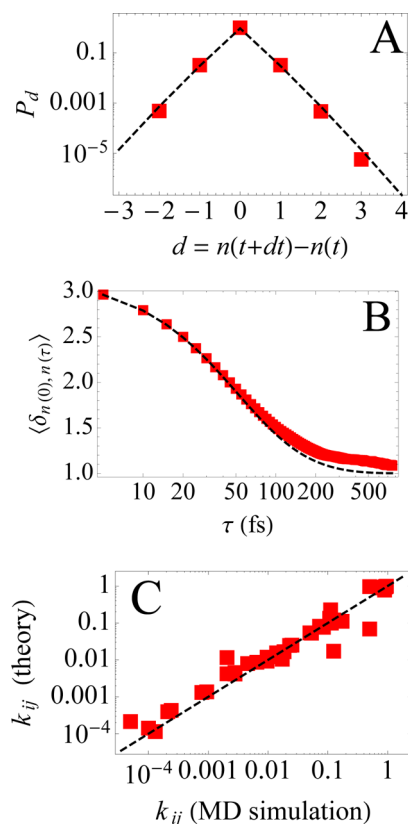


Figure 3. Panel A: The probability P_d of jump size estimated from the trajectory derived from MD simulation (red squares) is compared to the one predicted using the transition probabilities of the Markov process (dashed black line). Panel B: The normalized occupancy autocorrelation $\langle \delta_{n(0), n(\tau)} \rangle$ as estimated from the MD trajectory and as predicted from the transition probabilities of the Markov process. Panel C: We directly compare the transition probabilities k_{ij} for the probability of transition $i \rightarrow j$ empirically obtained from the MD trajectory to the ones predicted by using eq 23.

reproduced with high accuracy even though we did not utilize any information about it when inferring the transition probabilities of the Markov process. Moreover, our modeling also accurately captures the individual transition probabilities $\{k_{ij}\}$ over 4 orders of magnitude.

DISCUSSION AND SUMMARY

We have presented here a variational approach that computes $N \times N$ microscopic transition probabilities of a Markov process, given only knowledge of a stationary state population distribution and one trajectory-averaged dynamical property. In this approach, we maximize the path entropy subject to constraints; that is, we maximize the Caliber. We show that this method gives correct values of dynamical quantities in an example of molecular dynamics simulations of a water solvation shell around a single water molecule. This method may be useful for analyzing single-molecule experiments such as on ion channels,⁴ dynamics of neuron firing,⁷ and the dynamics of protein-sequence evolution,⁶ for example.

APPENDIX I: MD SIMULATION

We performed a molecular dynamics simulation on 233 water molecules^{25,26} at 300 K and at a constant volume using NAMD²⁷ with help of the Langevin thermostat. The oxygen atom of one of the water molecules was fixed at the origin. The

time step of integration was 1 fs and the trajectory was stored every 5 fs. Sampling the trajectory every 5 fs ensures that correlations in $n(t)$ have not vanished.

AUTHOR INFORMATION

Corresponding Author

*Email: pd2447@columbia.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

K.D. thanks grant 5R01GM090205-02 from the National Institute of General Medical Sciences. P.D. thanks Mr. Manas Rachh and Mr. Karthik Shekhar for numerous discussions about the topic and a critical reading of the manuscript. We would like to thank Kingshuk Ghosh and Steve Presse for helpful discussions on the manuscript.

REFERENCES

- (1) Weikl, T. R.; Palassini, M.; Dill, K. A. *Protein Sci.* **2004**, *13*, 822–829.
- (2) Lane, T. J.; Schwantes, C. R.; Beauchamp, K. A.; Pande, V. S. *J. Chem. Phys.* **2013**, *139*, 145104.
- (3) Lane, T. J.; Bowman, G. R.; Beauchamp, K.; Voelz, V. A.; Pande, V. S. *J. Am. Chem. Soc.* **2011**, *133*, 18413–18419.
- (4) Hille, B. *Ion Channels of Excitable Membranes*; Sinauer: Sunderland, MA, 2001; Vol. 507.
- (5) Kolomeisky, A. B.; Fisher, M. E. *Annu. Rev. Phys. Chem.* **2007**, *58*, 675–695.
- (6) Shekhar, K.; Ruberman, C. F.; Ferguson, A. L.; Barton, J. P.; Kardar, M.; Chakraborty, A. K. *Phys. Rev. E* **2013**, *88*, 062705.
- (7) Schneidman, E.; Berry, M. J.; Segev, R.; Bialek, W. *Nature* **2006**, *440*, 1007–1012.
- (8) Paulsson, J. *Phys. Life Rev.* **2005**, *2*, 157–175.
- (9) Dixit, P. D. *Biophys. J.* **2013**, *104*, 2743–2750.
- (10) Pressé, S.; Ghosh, K.; Lee, J.; Dill, K. A. *Rev. Mod. Phys.* **2013**, *85*, 1115.
- (11) Dixit, P. D. *J. Chem. Phys.* **2013**, *138*, 184111.
- (12) Crooks, G. E. *Phys. Rev. E* **2008**, *75*, 041119.
- (13) Jaynes, E. T. *Annu. Rev. Phys. Chem.* **1980**, *31*, 579–601.
- (14) Pressé, S.; Ghosh, K.; Lee, J.; Dill, K. A. *Rev. Mod. Phys.* **2013**, *85*, 1115–1141.
- (15) Stock, G.; Ghosh, K.; Dill, K. A. *J. Chem. Phys.* **2008**, *128*, 194102.
- (16) Wang, Q. A. *Chaos, Solitons, & Fractals* **2005**, *26*, 1045–1052.
- (17) Seitaridou, E.; Inamdar, M. M.; Phillips, R.; Ghosh, K.; Dill, K. J. *Phys. Chem. B* **2007**, *111*, 2288–2292.
- (18) Otten, M.; Stock, G. *J. Chem. Phys.* **2010**, *133*, 034119.
- (19) Ghosh, K.; Dill, K. A.; Inamdar, M. M.; Seitaridou, E.; Phillips, R. *Am. J. Phys.* **2006**, *74*, 123–133.
- (20) Filyukov, A.; Karpov, V. Y. *J. Eng. Phys. Thermophys.* **1967**, *13*, 416–419.
- (21) Cover, T. M.; Thomas, J. A. *Elements of Information Theory*; John Wiley & Sons: New York, 2012.
- (22) Caticha, A. *arXiv preprint arXiv:1212.6967*, **2012**.
- (23) Asthagiri, D.; Dixit, P. D.; Merchant, S.; Paulaitis, M. E.; Pratt, L. R.; Rempe, S. B.; Varma, S. *Chem. Phys. Lett.* **2010**, *485*, 1–7.
- (24) Merchant, S.; Asthagiri, D. *J. Chem. Phys.* **2009**, *130*, 195102.
- (25) Jorgensen, W.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (26) Neria, E.; Fischer, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 1902–1921.
- (27) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.