JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Inferring Transition Rates of Networks from Populations in Continuous-Time Markov Processes

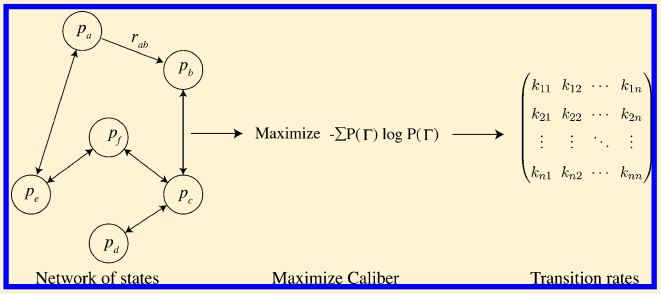
Purushottam D. Dixit,*[,†] Abhinav Jain,[‡] Gerhard Stock,[‡] and Ken A. Dill[§]

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

[‡]Institute of Physics and Freiburg Institute for Advanced Studies (FRIAS), Albert Ludwigs University, 79104 Freiburg, Germany

[§]Laufer Center for Quantitative Biology, Department of Chemistry, and Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11790, United States

ABSTRACT: We are interested inferring rate processes on networks. In particular, given a network's topology, the stationary populations on its nodes, and a few global dynamical observables, can we infer all the transition rates between nodes? We draw inferences using the principle of maximum caliber (maximum path entropy). We have previously derived results for discrete-time Markov processes. Here, we treat continuous-time processes, such as dynamics among metastable states of proteins. The present work leads to a particularly important analytical result: namely, that when the network is constrained only by a mean jump rate, the rate matrix is given by a square-root dependence of the rate, $k_{ab} \propto (\pi_b/\pi_a)^{1/2}$, on π_a and π_b, the stationary-state populations at nodes a and b. This leads to a fast way to estimate all of the microscopic rates in the system. As an illustration, we show that the method accurately predicts the nonequilibrium transition rates in an *in silico* gene expression network and transition probabilities among the metastable states of a small peptide at equilibrium. We note also that the method makes sensible predictions for so-called *extra-thermodynamic relationships*, such as those of Bronsted, Hammond, and others.



INTRODUCTION

We are interested in an inference problem in network science. Given the topology of a network and stationary populations at the nodes, what is the best model that can infer the rates of the dynamical flows along the edges? Here are examples. First, consider a spin model with a known stationary distribution, for example, those used in neuroscience,[1] protein evolution,[2] or colloidal sciences.[3] It is of great interest to infer the best dynamical process that is consistent with a given rate of spin flip while being consistent with the stationary distribution. Second, in systems biology, we often know the topology of a network of metabolites, or proteins, or regulatory elements. In addition, "-omics" experiments can estimate the abundances of the many metabolites or proteins or regulatory elements at the nodes during the steady-state functioning of a cell. However, this information alone is not sufficient to explain cell function. We also need to know the forward and backward rates of fluxes k_{ab} and k_{ba} between all nodes a and b, for example in metabolic networks.[4] Measuring all of these rates is practically impossible at present, particularly for large networks. Third, in structural biology, it is common to perform computer simulations of the conformations of biomolecules and infer Markov models among metastable states from those simulations.[5] Here, computing the populations of the states can be done rapidly, whereas computing the kinetic barriers between them is much slower.

For inferring *static properties* of distribution functions from a small number of observables, a standard approach is maximum entropy.[6−8] It has been used for a broad range of applications, including the flocking of birds,[9] variability in protein sequences,[10,11] firing rates of neurons,[1] dynamics of small systems,[12,13] and cell to cell variability in gene expression.[14] A related approach, called maximum caliber or maximum path entropy, is used for inferring *dynamical distribution functions*.[7,15−19] It is a general and rigorous method that seeks the state of maximized path entropy, rather than of maximized state entropy. It is this principle of maximum caliber (MaxCal) that we apply here to the problem of inferring the many transition rates between nodes on a network, given limited information.

In a previous work[18] we used MaxCal to derive transition probabilities p_{ab} of a *discrete-time* Markov chain. We imposed two types of constraints. First, we required the Markov chain to be consistent with a known stationary distribution over its states. Second, we required it to reproduce path-ensemble averages of few dynamical quantities. The discrete-time approach however is unsuitable in the case when the underlying stochastic process is continuous or approximately continuous in time. For example, intracellular biochemical reactions occur in continuous time. But, current experimental techniques allow simultaneous observation of multiple species at the time scale of seconds or minutes (see for example ref 20). Another example is long molecular dynamics (MD) simulations to study protein-wide motions.[5] The smallest

Received: June 8, 2015
Published: September 23, 2015

DOI: 10.1021/acs.jctc.5b00537
J. Chem. Theory Comput. 2015, 11, 5464−5472

time scale, the integration time step, in such calculations is usually of the order of 10^{-15} s while the time scale of protein-wide motions is usually of the order of 10^{-9}–10^{-6} s.[21] In both these cases and others, regardless of the discrete-time nature of the experimental observations, a continuous-time description of the underlying dynamics is required. In this work, we derive a continuous-time Markov jump process as a limiting case of the discrete-time maximum path entropy Markov chain.

We find that in the most general case the transition rates depend on the entire network of states, but for microscopically reversible processes they are determined locally and depend on the square roots of stationary probabilities. We illustrate the predictive power of our approach by studying (a) the nonequilibrium dynamics of an *in silico* gene expression network and (b) the dynamics of the network of metastable states of a small peptide.

THEORY

We first review our previous work on discrete-time Marko chains and then take the continuous-time limit.

Discrete-Time Markov Chain. We consider a discrete-time irreducible and aperiodic Markov chain on a network G with nodes V and edges E. We assume that $(a,a) \in E\ \forall\ a$. The unique stationary distribution over states is denoted by $\{\pi_a\}$. Markov property implies that the instantaneous probability $q_b(t + \delta t)$ of the chain being at any node b at time $t + \delta t$ depends only on the probabilities $q_a(t)$ at the previous time t and the *time homogeneous* transition probabilities p_{ab},

$$q_b(t + \delta t) = \sum_a q_a(t) p_{ab} \tag{1}$$

We assume that $p_{ab} \neq 0$ if $(a,b) \in E$ and zero otherwise.

The average of any arbitrary function r^i_{ab} that depends on two states a and b over the ensemble $\{\Gamma\}$ of stationary-state paths $\Gamma \equiv \ldots \rightarrow a \rightarrow b \rightarrow c \rightarrow \ldots$ of a fixed duration T is[18,22]

$$\langle r^i \rangle = \frac{1}{T} \sum_\Gamma p(\Gamma)(\ldots + r_{ab} + r_{bc} + \ldots) \tag{2}$$

$$= \sum_{a,b} \pi_a p_{ab} r^i_{ab} \tag{3}$$

Examples of dynamical variables include average flow of heat or mass and the distance traveled by a particle diffusing on a landscape. Clearly, there are infinitely many Markov chains that are consistent with the stationary distribution $\{\pi_a\}$ and a few path-ensemble averages $\langle r^i \rangle$. We recently showed that among these, the one that maximizes the path entropy

$$\mathcal{S} = -\sum_{a,b} \pi_a p_{ab} \log p_{ab} \tag{4}$$

is given by transition probabilities

$$p_{ab} = \frac{\lambda_a \beta_b}{\pi_a} \mathbf{W}_{ab} \tag{5}$$

where

$$\mathbf{W}_{ab} = \exp\left(-\sum_i \gamma_i r^i_{ab}\right) \tag{6}$$

when $(a,b) \in E$ and zero otherwise. γ_i are Lagrange multipliers associated with constraint $\langle r^i \rangle$ respectively that need to be adjusted simultaneously to reproduce the dynamical averages $\langle r^i \rangle$. Constants $\{\lambda_a\}$ and $\{\beta_a\}$ are determined by simultaneously solving

$$\sum_b \mathbf{W}_{ab} \lambda_b = \frac{\pi_a}{\beta_a}, \quad \sum_a \mathbf{W}_{ab} \beta_a = \frac{\pi_b}{\lambda_b} \tag{7}$$

Taking the Continuous-Time Limit. Here, we show how to take the continuous-time limit $\delta t \rightarrow 0$ in eq 5 to arrive at a Markov jump process. First, we assume that $r^i_{aa} = 0\ \forall\ a$ and i. This apparent restriction does not limit our choice of the dynamical constraint and is a mere notational convenience (see the Appendix for details). Second, in addition to $\langle r^i \rangle$, we also constrain the mean jump rate $\langle N \rangle$ where $N_{ab} = 1$ if $(a,b) \in E$ and $a \neq b$, $N_{ab} = 0$ otherwise. Note that $\langle N \rangle$ simply counts the total number of jumps per unit time. The matrix $\mathbf{W}$ in eq 6 becomes

$$\mathbf{W}_{ab} = \mathrm{e}^{-\alpha} \exp\left(-\sum_i \gamma_i r^i_{ab}\right) \tag{8}$$

when $(a,b) \in E$ and $a \neq b$, $\mathbf{W}_{aa} = 1$, and $\mathbf{W}_{ab} = 0$ otherwise. Here, α is the Lagrange multiplier associated with $\langle N \rangle$. Note that as α increases and $\langle N \rangle$ decreases. Since $\mathbf{W}_{aa} = 1$, we write

$$\mathbf{W} = \mathcal{I} + \mu \delta t \times \mathbf{\Delta} \tag{9}$$

where $\mathcal{I}$ is the identity matrix. We have identified $\mathrm{e}^{-\alpha} = \mu \delta t$. As we will see later, μ can be identified with the overall time scale of the Markov process. The transition probabilities are

$$p_{ab} = \mu \delta t \frac{\lambda_a \beta_b}{\pi_a} \mathbf{\Delta}_{ab} \tag{10}$$

when $(a,b) \in E$ and $a \neq b$,

$$p_{aa} = \frac{\lambda_a \beta_a}{\pi_a} \tag{11}$$

and $p_{ab} = 0$ otherwise.

The transition probabilities in eq 10 and eq 11 are amenable to a continuous-time limit $\delta t \rightarrow 0$ (or equivalently $\alpha \rightarrow \infty$ or $\langle N \rangle \rightarrow 0$) as long as the constants λ_a and β_a are well behaved in that limit (which they are; see Appendix for details). The transition rates k_{ab} of a continuous-time Markov jump process can be defined for $(a,b) \in E$ and $a \neq b$ as[23,24]

$$k_{ab} = \lim_{\delta t \rightarrow 0} \left(\frac{p_{ab}}{\delta t} \right) \tag{12}$$

The limiting procedure results in transition rates k_{ab} (see Appendix for details)

$$k_{ab} = \mu \sqrt{\frac{f_b g_a}{f_a g_b}} \sqrt{\frac{\pi_b}{\pi_a}} \mathbf{\Delta}_{ab} \tag{13}$$

if $(a,b) \in E$ and $a \neq b$ and $k_{aa} = -\sum_b k_{ab}$; $k_{ab} = 0$ otherwise. Equation 13 is the most general result of this work. The constants f_a and g_a are determined from *global* self-consistent equations:

$$f_a = \sum_b \mathbf{\Delta}_{ab} \sqrt{\frac{f_b \pi_b}{g_b}}, \quad g_a = \sum_b \mathbf{\Delta}_{ba} \sqrt{\frac{g_b \pi_b}{f_b}} \tag{14}$$

The transition rate k_{ab} between any two states a and b depends on the global structure of the network of connected states of the system.

Entropy Production and Detailed Balance. Detailed balance or microscopic reversibility is an important property of stochastic processes at thermodynamic equilibrium. Detailed

DOI: 10.1021/acs.jctc.5b00537
J. Chem. Theory Comput. 2015, 11, 5464−5472

balance stems from time reversal symmetry of the underlying dynamics. The entropy production rate $\dot{s}$ quantifies departure from microscopic reversibility.[25] $\dot{s}$ has a simple form for the maximum path entropy process. We have

$$\dot{s} = \sum_{a,b} \pi_a k_{ab} \log\left(\frac{k_{ab}}{k_{ba}}\right) = -\sum_i \gamma_i \left[\sum_{a,b} \pi_a k_{ab} (r^i_{ab} - r^i_{ba})\right] \tag{15}$$

In eq 15, constraints r^i_{ab} that are symmetric in a and b do not contribute to entropy production. Consequently, the entropy production is zero if all dynamical constraints are symmetric and the Markov process is automatically detail balanced. Indeed, if one wishes to impose detailed balance $\pi_a p_{ab} = \pi_b p_{ba}$ (equivalently, $\pi_a k_{ab} = \pi_b k_{ba}$) when the choice of constraint is asymmetric under time reversal, the inference problem becomes equivalent to constraining symmetrized forms of the constraints ${r_{ab}}^{i\dagger} = (1/2)(r^i_{ab} + r^i_{ba})$ (see ref 18 for a proof).

In this case, we have $\mathbf{\Delta} = \mathbf{\Delta}^{\mathrm{T}}$ and $\overline{f} = \overline{g}$. Thus

$$k_{ab} = \mu \sqrt{\frac{\pi_b}{\pi_a}}\, \mathbf{\Delta}_{ab} \tag{16}$$

if $(a,b) \in E$ and $a \neq b$. It is easy to check that eq 16 satisfies detailed balance with respect to the stationary distribution $\{\pi_a\}$. Interestingly, the transition rates of a detailed balanced Markov process are determined locally, i.e., by the properties of states a and b alone, and do not depend on the global structure of the network. In contrast, the same transition rates depend on the structure of the entire network when detailed balance is not satisfied (see eq 13). We note that the square-root dependence of rates on the stationary distribution was previously anticipated by studying the discretized version of a reversible Smoluchowski equation describing a particle diffusing in one dimension.[26,27]

Equation 16 takes a particularly simple form when we do not constrain any dynamical quantity apart from $\langle N \rangle$, or the mean jump rate. In other words, $r^i_{ab} = 0$ for all a,b, and i. We have

$$k_{ab} = \mu \sqrt{\frac{\pi_b}{\pi_a}} \tag{17}$$

if $(a,b) \in E$ and $a \neq b$. This result should be contrasted with other popular functional forms of the transition rates that satisfy a prescribed stationary distribution, for example, Glauber dynamics,[28]

$$k^{\mathrm{glauber}}_{ab} = \mu \frac{\pi_b}{\pi_a + \pi_b} \tag{18}$$

and Metropolis dynamics,[29]

$$k^{\mathrm{metropolis}}_{ab} = \mu \min\left(1, \frac{\pi_b}{\pi_a}\right) \tag{19}$$

Historically, Glauber developed his dynamics to study magnetic spins. The particular form of the transition rates was motivated by "a desire for simplicity".[28] The derivation of square-root dynamics presented here, while not resorting to an *ad hoc* prescription for transition rates, follows the same intuitive principle, i.e., finding the simplest model that is consistent with a given set stationary distribution and an overall spin flip rate and surprisingly predicts a dynamics that is *qualitatively* different from Glauber and Metropolis dynamics.

Comment on Chemical Reaction Rates. In organic chemistry, so-called *extra-thermodynamic relationships* express empirical observations about how the rates and mechanisms of certain types of chemical reactions are related to their equilibria. These go by the name of the Bronsted relation (often applied to acid−base catalysis),[30] the Polanyi relationship for surface catalysis,[31] Marcus theory for electron transfer reactions in solution,[32] or Φ-value analysis for protein folding.[33] In short, these approaches express that rates are related to equilibrium constants in the form of

$$k(x) = cK(x)^{\alpha} \tag{20}$$

where x is a variable representing a systematic change, such as a series of different acids of different pK_a's, $K(x)$ are the equilibrium constants for that series, and $k(x)$ are the rates of reaction. α is a value that usually ranges from 0 to 1, expressing the degree of resemblance of the transition state to the products of the reaction.

To put the preceding observations in the perspective of the square-root dynamics, consider a two-state system where state a is the ensemble of reactant states and b is the ensemble of product states such that $\pi_b > \pi_a$. We have neglected the population in the transition state since it is much smaller than the state populations at equilibrium. We have equilibrium constant $K = \pi_b/\pi_a > 1$. If no dynamical constraints are imposed, the transition rate k_{ab} of the $a \to b$ reaction according to the maximum path entropy process is given by $k_{ab} = \mu(\pi_b/\pi_a)^{1/2} = \mu K^{1/2}$. In short, our maximum path entropy approach predicts that $\alpha = 1/2$ in eq 20 is the most parsimonious assumption for the reaction mechanism prior to any knowledge of global rate information, implying that the transition state is halfway between reactants and products.

How do we rationalize $\alpha \neq 1/2$ from a maximum path entropy perspective? We recognize the equilibrium probabilities as $\pi_a \propto e^{-\beta G_a}$ and $\pi_b \propto e^{-\beta G_b}$, respectively, where G_a and G_b are the free energies of the two states such that $G_b < G_a$. Here, β is the inverse temperature. Observe that $\Delta H = |G_b - G_a| = (1/\beta)\log(\pi_b/\pi_a) = (1/\beta)\log K$ is the average the the amount of heat released to the surrounding bath when the ensemble of reactant states a is converted into the ensemble of product states b. The magnitude of the heat transfer for $a \to b$ and $b \to a$ transition is the same and equal to ΔH but the sign is opposite. We ask the following question. If, along with the stationary distribution, we constrain $\langle \Delta H \rangle$, the average energy exchanged with the thermal bath per unit time, we predict a Markov process

$$k_{ab} = \mu \sqrt{\frac{\pi_b}{\pi_a}}\, e^{-\gamma \Delta H} \tag{21}$$

$$\Rightarrow \log k_{ab} = \alpha \log K + c \tag{22}$$

where

$$\alpha = \frac{1}{2} - \frac{\gamma}{\beta} \tag{23}$$

and $c = \log \mu$. Here, γ is the Lagrange multiplier associated with the rate of heat exchange with the bath. We hope that the maximum path entropy basis of extra-thermodynamic relationships will be explored in detail in later studies.

DOI: 10.1021/acs.jctc.5b00537
J. Chem. Theory Comput. 2015, 11, 5464−5472

ILLUSTRATIONS

To illustrate the predictive power of the developed framework, we study a nonequilibrium *in silico* gene expression network and the equilibrium dynamics of a small peptide.

Gene Expression Network. Here, we show how eq 13 can accurately estimate a nonequilibrium chemical master equation (CME) from stationary distributions and global dynamical averages. Consider a biochemical circuit where two genes are adjacent to a constitutively expressing promoter region (Figure 1). We assume that the genes are either transcribed individually

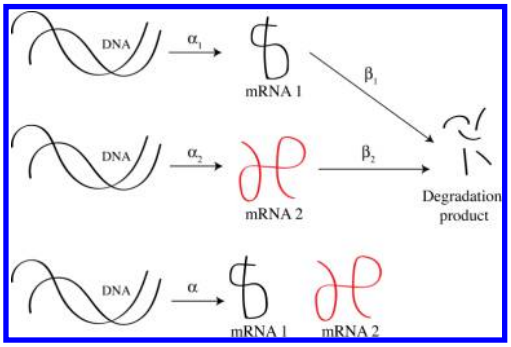


Figure 1. Schematic of two genes that are simultaneously transcribed by the same promoter. Rates α_1, α_2, and α represent the synthesis rates and rates β_1 and β_2 represent the degradation rates.

or simultaneously and that they are degraded individually. In this toy example, the copy numbers of the two mRNA molecules are stochastically coupled. We first construct a CME to mimic the biochemical circuit *in silico*. There are five rate parameters in the CME: three synthesis rates α_1, α_2, and α and two degradation rates β_1 and β_2. If n_1 and n_2 are the number of molecules of the first and the second mRNA, the chemical master equation describing the system has the following form

$$\begin{aligned}\frac{dp(n_1, n_2; t)}{dt} &= \alpha_1(p(n_1 - 1, n_2) - p(n_1, n_2)) \\ &+ \alpha_2(p(n_1, n_2 - 1) - p(n_1, n_2)) \\ &+ \alpha(p(n_1 - 1, n_2 - 1) - p(n_1, n_2)) \\ &+ \beta_1((n_1 + 1)p(n_1 + 1, n_2) - n_1p(n_1, n_2)) \\ &+ \beta_2((n_2 + 1)p(n_1, n_2 + 1) - n_2p(n_1, n_2))\end{aligned} \tag{24}$$

Here, $p(n_1,n_2;t)$ is the instantaneous probability of having n_1 and n_2 molecules of mRNA 1 and 2, respectively, at time t. The terms correspond to individual syntheses of mRNA 1 and mRNA 2, the simultaneous synthesis of both mRNAs, and the degradation of mRNA 1 and mRNA 2.

We assume that we can experimentally estimate the joint probability distribution $p_{ss}(n_1,n_2)$ of the mRNA copy numbers at steady state. Additionally, we also assume that we have two reporters that count the total number of expression events and the total number of degradation events, respectively. Note that the reporters are agnostic to which of the two RNAs has been synthesized or degraded. With only these three pieces of information and the underlying network of possible transitions (only one gene can be degraded and up to two genes can be synthesized at any particular instant), we ask whether we can estimate the entire transition rate matrix of the system.

We choose the following parameters for the CME: (α_1, α_2, α, β_1, β_2) = (1, 0.5, 2.5, 5, 10). The choice ensures that the number of any of the two mRNA molecules is limited to <6 which results in a small system size with the total number of states = 6 × 6 = 36. In Figure 2 we show the numerically observed joint distribution $p_{ss}(n_1,n_2)$. The correlated pattern of expression is apparent; the conditional probability $p_{ss}(n_2|n_1)$ depends on n_1. The higher the value of n_1, the higher n_2 values become more probable as a result of the correlated expression.

In the CME, while there are only five rate parameters, there are 145 possible transitions. There are 30 + 30 transitions that correspond to synthesis of mRNA 1 or 2. There are 30 + 30 transitions that correspond to degradation of mRNA 1 or 2, and there are 25 transitions that correspond to simultaneous synthesis of both mRNAs. Each transition has its own rate constant, but the rates are related to each other. For example, the transition (2, 3) → (2,2) has a rate constant $3\beta_2$ and the transition (2,4) → (2,3) has a rate constant $4\beta_2$. Moreover, many other transition rates are equal to each other, for example, the rates for transitions that increase the first mRNA copy number such as (1,2) → (2,2) and (3,1) → (4,1) are equal to α_1 and so on.

Even though the transition rates of the original CME are intricately related to each other, the developed inference procedure is agnostic to that detail. The MaxCal inference treats all allowed transition rates as potentially different entities. Nonetheless, we find that eq 13 predicts the full 145 rate parameters, without the knowledge of the differential rates of synthesis and degradation of the two mRNAs with sufficient

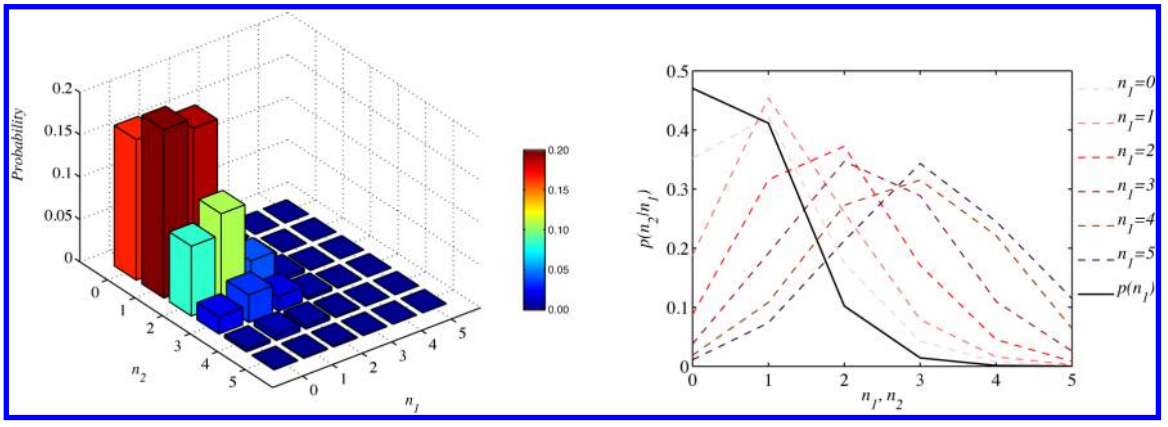


Figure 2. Left: Joint probability distribution $p_{ss}(n_1,n_2)$ of the two mRNA copy numbers. Right: Conditional probability $p_{ss}(n_2|n_1)$ at different values of n_1 in the stationary state.

DOI: 10.1021/acs.jctc.5b00537
J. Chem. Theory Comput. 2015, 11, 5464−5472

accuracy (see Appendix for details of the fitting procedure and see Figure 3). Taken to larger scale, it implies that using steady-state data on cell-to-cell variability in gene expression and a few overall kinetic measurements, we can obtain a full set of chemical master equations and infer regulatory details. This inferred model is optimal in the maximum caliber sense.

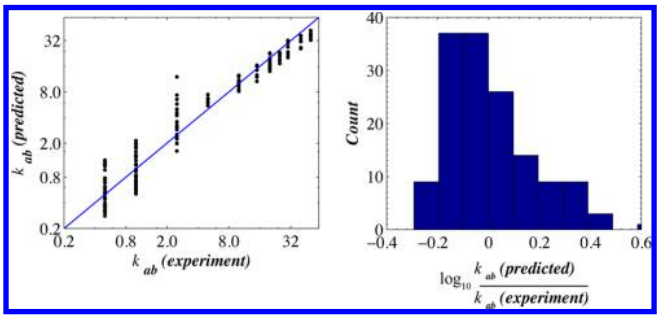


Figure 3. Left: Predicted rates vs toy model values for all 145 possible transitions in the gene expression model. Right: Histogram of logarithmic error between predicted and experimental rates.

Equilibrium Dynamics of a Peptide. Next, we study the equilibrium dynamics of metastable states of a small peptide. A metastable state is an ensemble of geometrically proximal microstructures that have a significant aggregate population. On the one hand, the metastable states of moderately sized proteins can be characterized relatively cheaply and in a highly parallelizable manner using a variety of techniques including but not limited to replica exchange molecular dynamics (REMD),[34] accelerated molecular dynamics,[35] umbrella sampling,[36] and metadynamics.[37] On the other hand, transitions rates between these metastable states are difficult to estimate and may require long nonparallelizable molecular dynamics simulations.[5]

We show that, with appropriate constraints, the framework developed here can take as input the ensemble of metastable states and an overall rate measurement (for example, a folding rate or a relaxation rate) from experiments and predict transition rates between all metastable states. We put to test our framework using a previous extensive molecular dynamics simulation of a seven alanine residue peptide.[38] The simulation led to identification of 32 metastable ensembles of the peptide along with their equilibrium probabilities $\{\pi_a\}$[39] (see Figure 4). From trajectories extracted from the simulation, we also estimated symmetrized transition probabilities p_{ab} (*not* transition rates k_{ab}) as the fraction of the events in the long trajectory when the peptide was in state a at time t and transitioned into state b at time $t + \delta t$ ($\delta t = 1$ ps in a MD simulation of total duration $T = 800$ ns) (see Appendix for details). While a 800 ns long simulation is not computationally cheap, it allows us a direct comparison between rates observed in the MD simulation and our corresponding predictions. Moreover, as stated earlier, the ensembles of the 32 metastable states and their relative probabilities can also be obtained using a number of efficient numerical techniques.

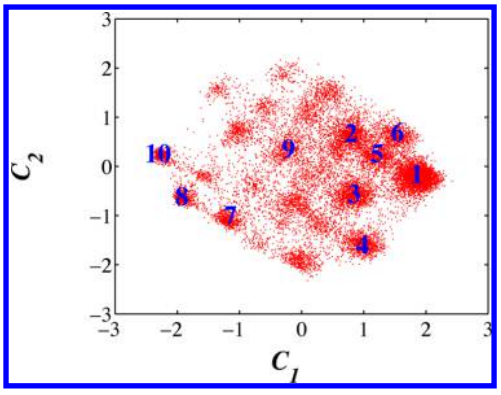


Figure 4. Two principal coordinates from molecular dynamics simulations of the alanine-7 peptide,[38] showing that the microscopic structures can be lumped into well separated metastable states. The 10 most populated states are labeled in decreasing order of state probability.

To predict the observed transition probabilities p_{ab}, we first need to estimate the continuous-time transition rates k_{ab} using eq 16. First, we need to identify a suitable dynamical constraint. The accuracy of our predictions depends on how well the dynamical constraint captures the diffusion of the peptide on the free energy landscape. The dependence of model predictions on the choice of the constraints is a common aspect of maximum-entropy methods. See refs 7, 12, 14, 40, and 41 for a discussion on the role of constraints in maximum entropy methods.

In order to arrive at the form of the dynamical constraint, we make two observations: the transition rate between two states should decrease (1) when the average conformational separation is increased, keeping the stationary probabilities and the free energy barrier constant, and (2) when the free energy barrier is increased, keeping the average conformational distance and the stationary-state probabilities constant. In this study, we neglected the effect of energetic barriers and focused on geometric separation. We compared three different dynamical constraints that quantify geometric separation between metastable-state ensembles in Figure 5.

In panel A, information about the geometric separation was not included and we only constrained the mean jump rate and used eq 17. In panel B, we used as a constraint the distance traveled by the protein in internal coordinate representation;

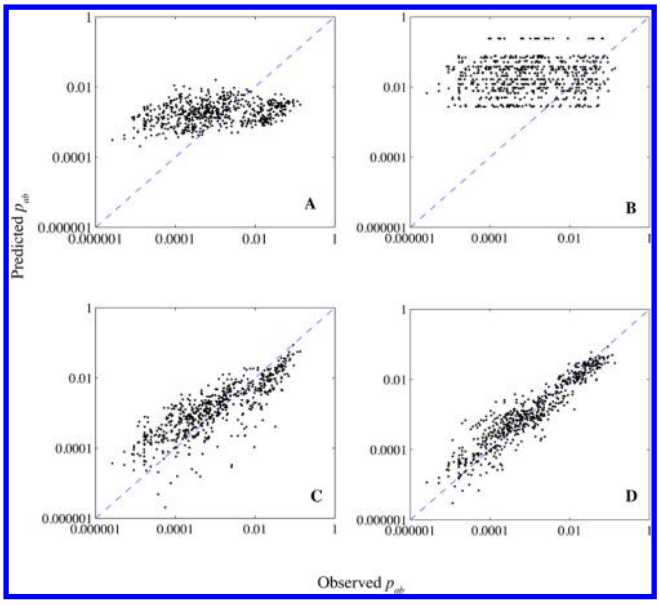


Figure 5. Predicton of transition probabilities p_{ab} for four different dynamical constraints. (A) Mean jump rate. (B) Distance between average structures of the states. (C) Symmetrized Kullback−Leibler divergence between the states. (D) Average ensemble distance between the states.

DOI: 10.1021/acs.jctc.5b00537
J. Chem. Theory Comput. 2015, 11, 5464−5472

calculated between the mean structures of each of the 32 metastable states, $r_{ab} = |x_a - x_b|$ where x_a are internal coordinates of the average structure of metastable state a. In panel C, we used as a constraint the symmetrized Kullback−Leibler divergence between states estimated by assuming a multinormal distribution around the mean structure x_a in the internal coordinate space, $r_{ab} = (KL(a,b) + KL(b,a))/2$. Finally, in panel D, we choose an ensemble dependent geometric constraint. For any two metastable states a and b and microstates v and w such that $v \in a$ and $w \in b$, we define $r_{ab} = (\int_{v\in a, w\in b} \|v - w\|_2 \, dv \, dw)^{1/2}$. The distance between any two microstates v and w is defined as the Euclidean distance between their internal coordinate representation. See Appendix for details of the calculation.

Using eq 16, the numerically estimated stationary probabilities, and the chosen geometric constraint, we arrive at the functional form of the transition rates k_{ab}. The predicted 32×32 transition rate matrix had only two free parameters; the time scale μ and the Lagrange multiplier γ associated with $\langle r \rangle$. In order to estimate the transition probabilities p_{ab} from the predicted transition rates k_{ab}, we need the discretization time scale δt, a third parameter. At a given discretization time scale, the transition probability matrix $\mathbf{p}$ is related to the transition rate matrix $\mathbf{K}$ as follows

$$\mathbf{p} = e^{\delta t \mathbf{K}} \tag{25}$$

From eq 25 and eq 16 it is clear that μ and δt can be combined together. Thus, we only needed to determine two parameters, γ and $\mu\delta t$. Note that the maximum path entropy approach guesses the parametric form of the transition probabilities; one can use any suitable numerical technique and experimental information to estimate the parameters. Since we have access to microscopic data, in this proof of principle study, we used the entire transition probability matrix to fit the two free parameters.

In Figure 5, we see that the accuracy of the prediction indeed depends on the nature of the constraint. At least for a small peptide, a Markov jump process constrained to reproduce the average microscopic distance traveled per unit time results in excellent agreement with full MD simulations (panel C). Inclusion of free energy barriers is likely to improve these estimates, for example, for larger proteins.

DISCUSSION

Here, we describe theory that takes a given network, steady-state populations on its nodes, and a few of global dynamical constraints and finds the microscopic transition rates among all of the nodes. We do this by maximizing the path entropy of a random walk on the network. A main finding is that the maximum path entropy transition rates are proportional to the square root of ratios of the state populations. We showed that we can predict transition rates among metastable rates of a peptide given only the knowledge of the states and an overall rate quantity. We believe this treatment could be useful in many areas of network modeling, including in spin-glass models of the immune system,[2,11] colloidal assemblies,[3] neuronal networks,[1] master-equation models of noisy gene expression,[14,42] and in the browsing behavior of web crawlers on the Internet.[43]

APPENDIX

Deriving the Maximum Path Entropy Transition Rates

We present a brief overview of the results in ref 18. We first derive the discrete-time maximum path entropy Markov chain and then take the continuous-time limit.

We know that the stationary distribution $\{\pi_a\}$ and the transition probabilities $\{p_{ab}\}$ are not independent of each other. We have

$$\sum_b \pi_a p_{ab} = \pi_a, \qquad \sum_a \pi_a p_{ab} = \pi_b \tag{26}$$

We maximize the path entropy $\mathcal{S}$ with respect to the transition probabilities while imposing constraints in eq 3 and eq 26. For notational simplicity, we derive eq 13 for one dynamical constraint r_{ab}. We have assumed that $r_{aa} = 0$ for all a. Generalization to multiple constraints is straightforward. Using the method of Lagrange multipliers, we write the unconstrained caliber C[7]

$$C = \mathcal{S} + \sum_a m_a\Big(\sum_b \pi_a p_{ab} - \pi_a\Big) + \sum_b l_b\Big(\sum_a \pi_a p_{ab} - \pi_b\Big) - \alpha\Big(\sum_{a,b} \pi_a p_{ab} N_{ab} - \langle N \rangle\Big) - \gamma\Big(\sum_{a,b} \pi_a p_{ab} r_{ab} - \langle r \rangle\Big) \tag{27}$$

Differentiating the caliber in eq 27 with respect to p_{ab} and setting the derivative to zero,

$$\pi_a(\log p_{ab} + 1) = m_a\pi_a + l_b\pi_a - \alpha N_{ab}\pi_a - \gamma r_{ab}\pi_a$$
$$\Rightarrow p_{ab} = \frac{\beta_a}{\pi_a}\lambda_b \mathbf{W}_{ab} \tag{28}$$

where $\beta_a/\pi_a = e^{m_a - 1}$, $\lambda_b = e^{l_b}$, and $\mathbf{W}_{ab} = e^{-\alpha N_{ab} - \gamma r_{ab}}$.

For a given value of α and γ, the modified Lagrange multipliers βs and λs are determined by self consistently solving eqs 26. We have

$$\sum_b p_{ab} = 1 \Rightarrow \sum_b \mathbf{W}_{ab}\lambda_b = \frac{\pi_a}{\beta_a}$$
$$\sum_a \pi_a p_{ab} = \pi_b \Rightarrow \sum_b \mathbf{W}_{ba}\beta_b = \frac{\pi_a}{\lambda_a} \tag{29}$$

To further simplify eqs 29, we identify $\bar{\lambda}$ and $\bar{\beta}$ as the vectors of Lagrange multipliers and define $\mathcal{D}[\bar{x}]_a = \pi_a/x_a$, a nonlinear operator on vectors $\bar{x}$. We have

$$\mathbf{W}\bar{\lambda} = \mathcal{D}[\bar{\beta}], \qquad \mathbf{W}^{\mathrm{T}}\bar{\beta} = \mathcal{D}[\bar{\lambda}] \tag{30}$$

$$\Rightarrow \mathcal{D}[\mathbf{W}\bar{\lambda}] = \bar{\beta}, \qquad \mathcal{D}[\mathbf{W}^{\mathrm{T}}\bar{\beta}] = \bar{\lambda} \tag{31}$$

Recognizing that $r_{aa} = 0$, we write $\mathbf{W} = \mathbf{I} + \mu\delta t\mathbf{\Delta}$ where $e^{-\gamma} = \mu\delta t$, $\mathbf{\Delta}_{ab} = 0$ if $(a,b) \notin E$, and $\mathbf{\Delta}_{ab} = e^{-\gamma r_{ab}}$ when $(a,b) \in E$ and $a \neq b$. We have

$$\frac{p_a}{\lambda_a + \mu\delta t f_a} = \beta_a \tag{32}$$

$$\frac{p_a}{\beta_a + \mu\delta t g_a} = \lambda_a \tag{33}$$

where

$$\bar{f} = \mathbf{\Delta}\bar{\lambda} \tag{34}$$

$$\bar{g} = \mathbf{\Delta}^{\mathrm{T}}\bar{\beta} \tag{35}$$

We solve these two equations algebraicly recognizing that f_a and g_a do not directly depend on λ_a and β_a. The positive roots are

DOI: 10.1021/acs.jctc.5b00537
J. Chem. Theory Comput. 2015, 11, 5464−5472

$$\lambda_a = \frac{\sqrt{f_a}\sqrt{g_a}\sqrt{\delta t^2 f_a g_a \mu^2 + 4\pi_a} - \delta t f_a g_a \mu}{2g_a} \tag{36}$$

$$\beta_a = \frac{\sqrt{f_a}\sqrt{g_a}\sqrt{\delta t^2 f_a g_a \mu^2 + 4\pi_a} - \delta t f_a g_a \mu}{2f_a} \tag{37}$$

Since we want to take the limit $\delta t \to 0$, we're only interested in transition probabilities p_{ab} up to the leading order in δt and $p_{ab} = \mu\delta t(\beta_a/\pi_a)\lambda_b\Delta_{ab}$ that already has a δt term; we only need to zero order terms for $\overline{\lambda}$ and $\overline{\beta}$ as $\delta t \to 0$. We solve the algebraic eqs 33 simultaneously for λ_a and β_a in terms of f_a, g_a, π_a, and δt. We then take only the zero order terms in δt. We have

$$\lambda_a = \sqrt{\frac{f_a \pi_a}{g_a}} \tag{38}$$

$$\beta_a = \sqrt{\frac{g_a \pi_a}{f_a}} \tag{39}$$

Equation 39 and eq 35 can be self-consistently solved for $\overline{\lambda}$ and $\overline{\beta}$, and we can obtain $\overline{f} = \mathbf{\Delta}\overline{\lambda}$ and $\overline{g} = \mathbf{\Delta}^{\mathrm{T}}\overline{\beta}$.

Now consider a continuous-time Markovian random walk on the same network G. The instantaneous probability of the walk being at a node b at time t, $q_b(t)$, is governed by the master equation

$$\frac{dq_b(t)}{\delta t} = \sum_a k_{ab} q_a(t) - \sum_a k_{ba} q_b(t) = \sum_a \mathbf{\Omega}_{ba} q_a(t) \tag{40}$$

where the time independent rate of transition k_{ab} from node a to node b is non-zero if and only if $(a, b) \in E$. If we know $q_a(t)$, symbolically, we can solve eq 40 for a later time $t + \delta t$,

$$q_b(t + \delta t) = \sum_a q_a(t)\, p_{ab}(\delta t) \tag{41}$$

where the transition probability matrix $\mathbf{p}$ is related to the transition rate matrix $\mathbf{\Omega}$ through

$$\mathbf{p} = e^{\mathbf{\Omega}\delta t} \approx \mathbf{I} + \mathbf{\Omega}\delta t \tag{42}$$

where the last approximation holds true only when $\delta t \to 0$. Comparing eq 1 and eq 42, we arrive at eq 13.

When Dynamical Constraints Are Such That $r_{aa} \neq 0$

As mentioned in the main text, we have assumed for convenience that the dynamical constraints r_{ab} are such that $r_{aa} = 0$. Here, we show that when $r_{aa} \neq 0$, the maximum entropy problem is equivalent to constraining modified constraints $r^{\dagger}_{ab} = r_{ab} - (1/2)(r_{aa} + r_{bb})$.

We start with recognizing $\mathbf{W} = \mathbf{J} + \mu\delta t\, \mathbf{\Delta}$ as discussed earlier, where $\mathbf{J}$ is a diagonal matrix such that $\mathbf{J}_{aa} = e^{-\gamma r_{aa}}$. We have

$$\frac{\pi_a}{j_a\lambda_a + \mu\delta t f_a} = \beta_a \tag{43}$$

$$\frac{\pi_a}{j_a\beta_a + \mu\delta t g_a} = \lambda_a \tag{44}$$

where $j_a = \mathbf{J}_{aa}$ and $\overline{f} = \mathbf{\Delta}\overline{\lambda}$ and $\overline{g} = \mathbf{\Delta}^{\mathrm{T}}\overline{\beta}$. Again solving to zero order in δt, we get

$$\lambda_a = \sqrt{\frac{f_a \pi_a}{g_a j_a}} \tag{45}$$

$$\beta_a = \sqrt{\frac{\pi_a g_a}{f_a j_a}} \tag{46}$$

Finally, the transition probability p_{ab} is given by

$$\begin{aligned} p_{ab} &= \mu\delta t \frac{\beta_a}{\pi_a}\lambda_b e^{-\gamma r_{ab}} = \mu\delta t \sqrt{\frac{\pi_b}{\pi_a}}\sqrt{\frac{f_b g_a}{f_a g_b}}\frac{1}{\sqrt{j_a j_b}} e^{-\gamma r_{ab}} \\ &= \mu\delta t \sqrt{\frac{\pi_b}{\pi_a}}\sqrt{\frac{f_b g_a}{f_a g_b}} e^{-(\gamma r_{ab} - (\gamma/2)(r_{aa} + r_{bb}))} \\ &= \mu\delta t \sqrt{\frac{\pi_b}{\pi_a}}\sqrt{\frac{f_b g_a}{f_a g_b}} e^{-\gamma r^{\dagger}_{ab}} \end{aligned} \tag{47}$$

where $r^{\dagger}_{ab} = r_{ab} - (1/2)(r_{aa} + r_{bb})$. Comparing eq 47 to eq 13, it is clear that the problem of constraining a dynamical constraint r_{ab} such that r_{aa} is nonzero is equivalent to constraining a modified dynamical constraint $r^{\dagger}_{ab} = r_{ab} - (1/2)(r_{aa} + r_{bb})$. Note that, by definition, $r^{\dagger}_{aa} = 0$. Thus, for convenience, we assume that this transformation is already performed and $r_{aa} = 0$.

Constructing and Fitting the Maximum Entropy Markov Proces for the Gene Network

From the numerical experiments, we obtain an accurate estimate of the stationary-state probability distribution $p_{ss}(n_1,n_2)$. Let $a \equiv (n_1, n_2)$ and $b \equiv (n_1^{\dagger}, n_2^{\dagger})$ be any two states of the system. We know that there is a directed edge from state a to state b, $(a,b) \in E$, *iff*

1. Synthesis or degradation of mRNA 1: $n_1 = n_1^{\dagger} \pm 1$ and $n_2 = n_2^{\dagger}$
2. Synthesis or degradation of mRNA 2: $n_1 = n_1^{\dagger}$ and $n_2 = n_2^{\dagger} \pm 1$
3. Simultaneous synthesis: $n_1^{\dagger} = n_1 + 1$ and $n_2^{\dagger} = n_2 + 1$

As mentioned in the main text, the "experiments" tell us the total number of degradation and synthesis events per unit time but do not have the ability to distinguish between the two mRNAs. We constrain two quantities, the number of degradation events per unit time and the number of synthesis events per unit time. Accordingly, we set $\mathbf{\Delta}$ in eq 13 as follows:

1. No edge between nodes a and b: $\mathbf{\Delta}_{ab} = 0$ if $(a, b) \notin E$
2. Synthesis events: $\mathbf{\Delta}_{ab} = \eta$ if $n_1 = n_1^{\dagger}$ and $n_2 + 1 = n_2^{\dagger}$ or $n_1 + 1 = n_1^{\dagger}$ and $n_2 = n_2^{\dagger}$ or $n_1 + 1 = n_1^{\dagger}$ and $n_2 + 1 = n_2^{\dagger}$
3. Degradation of mRNA 1: $\mathbf{\Delta}_{ab} = \zeta$ if $n_1 = n_1^{\dagger} + 1$ and $n_2 = n_2^{\dagger}$
4. Degradation of mRNA2: $\mathbf{\Delta}_{ab} = \zeta$ if $n_1 = n_1^{\dagger}$ and $n_2 = n_2^{\dagger} + 1$

Here, $\eta \geq 0$ and $\zeta \geq 0$ are exponentials of Lagrange multipliers similar to those used in the main text. The Lagrange rate constant μ in eq 13 is assumed to be absorbed in the Lagrange multipliers. The next step is to determine $\overline{f}$, $\overline{g}$, $\overline{\lambda}$, and $\overline{\beta}$ using numerically estimated $p_{ss}(n_1,n_2)$ and $\mathbf{\Delta}$. In order to determine the transition rate matrix $\mathbf{\Omega}$ for any value of η and ζ, we solve eq 35 and eq 39 self-consistently.

Given that we have access to all rate constants in this proof of principles work, we minimize the error between the known rate constants of eq 24 and those predicted by eq 13 by varying η and ζ using a simulated annealing protocol. We find that $\eta \approx 0.16$ and $\zeta \approx 82$ resulted in the best agreement between the known and the

DOI: 10.1021/acs.jctc.5b00537
J. Chem. Theory Comput. 2015, 11, 5464−5472

predicted rate constants. Multiple simulated annealing runs predicted rates that were identical to the ones shown in the main text. Note that although we used the entire transition matrix to learn the Lagrange multipliers, it is equally possible to learn them from global dynamical constraints.

Symmetrization of Observed Transition Probabilities

Once the metastable states have been identified, the stationary probabilities can be estimated as

$$\pi_a = \frac{n_a}{n_s} \tag{48}$$

where n_a is the number of samples in state a and n_s is the total number of samples. The transition probabilities can be estimated as

$$p_{ab} = \frac{1}{\pi_a} \frac{n_{ab}}{n_s - 1} \tag{49}$$

where n_{ab} is the number of times the system starts in state a and ends in state b in one time step. Since the data come from an equilibrium simulation, ideally we expect detailed balance; $p_{ab}\pi_a = p_{ba}\pi_b$ if the identification of metastable states has achieved time scale separation. In order to remove sampling errors, we use the symmetrized form of the transition probabilities

$$p_{ab}^{\mathrm{sym}} = \frac{\pi_a p_{ab} + \pi_b p_{ba}}{\pi_a} \tag{50}$$

Figure 6 shows the comparison between the symmetrized and the unsymmetrized estimate of p_{ab}. In this work, we used the symmetric form p_{ab}^{sym} in our analysis.

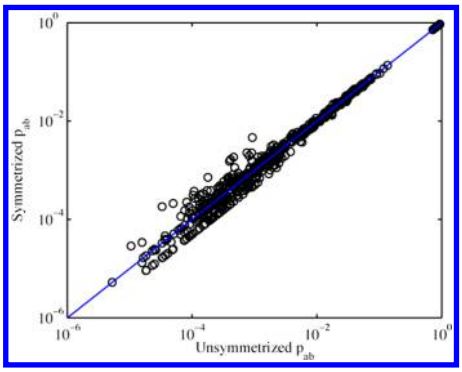


Figure 6. Comparison between symmetrized and unsymmetrized estimate of p_{ab}.

Estimating the Dynamical Constraint

We tested three different types of dynamical constraints. First, we estimated the coordinates x_a of the average structure of each of the 32 metastable states in internal coordinate representation. The dynamical constraint was simply $r_{ab} = |x_a - x_b|$. This dynamical constraint does not take into account the ensemble of microstates that belong to a metastable state. Next, we tested the symmetrized Kullback−Leibler divergence between two states as the dynamical constraint. The estimation of Kullback−Leibler divergence requires the knowledge of the underlying distribution. We assumed that the microstates are distributed multinormally around the mean structure and estimated the covariance matrix σ_a of each microstate. The Kullback−Leibler divergence was simply

$$r_{ab} = \frac{1}{4}\left(\mathrm{tr}(\sigma_a^{-1}\sigma_b) + (x_a - x_b)^{\mathrm{T}}\sigma_a^{-1}(x_a - x_b) - 32 + \log\left(\frac{|\sigma_a|}{|\sigma_b|}\right)\right) + \frac{1}{4}\left(\mathrm{tr}(\sigma_b^{-1}\sigma_a) + (x_b - x_a)^{\mathrm{T}}\sigma_b^{-1}(x_b - x_a) - 32 + \log\left(\frac{|\sigma_b|}{|\sigma_a|}\right)\right) \tag{51}$$

Finally, we estimated the average distance traveled by microstates as a dynamical constraint as explained in the main text. We have

$$r_{ab} = \left(\int_{v\in a, w\in b} \|v - w\|_2 \, \mathrm{d}v \, \mathrm{d}w\right)^{1/2} \tag{52}$$

AUTHOR INFORMATION

Corresponding Author

*E-mail: pd2447@cumc.columbia.edu.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Schneidman, E.; Berry, M. J.; Segev, R.; Bialek, W. *Nature* **2006**, *440*, 1007−1012.
(2) Shekhar, K.; Ruberman, C. F.; Ferguson, A. L.; Barton, J. P.; Kardar, M.; Chakraborty, A. K. *Phys. Rev. E* **2013**, *88*, 062705.
(3) Han, Y.; Shokef, Y.; Alsayed, A. M.; Yunker, P.; Lubensky, T. C.; Yodh, A. G. *Nature* **2008**, *456*, 898−903.
(4) Orth, J. D.; Thiele, I.; Palsson, B. Ø. *Nat. Biotechnol.* **2010**, *28*, 245−248.
(5) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. *Science* **2010**, *330*, 341−346.
(6) Jaynes, E. T. *Phys. Rev.* **1957**, *106*, 620.
(7) Pressé, S.; Ghosh, K.; Lee, J.; Dill, K. A. *Rev. Mod. Phys.* **2013**, *85*, 1115−1141.
(8) Shore, J.; Johnson, R. *IEEE Trans. Inf. Theory* **1980**, *26*, 26−37.
(9) Bialek, W.; Cavagna, A.; Giardina, I.; Mora, T.; Silvestri, E.; Viale, M.; Walczak, A. M. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 4786−4791.
(10) Weigt, M.; White, R. A.; Szurmant, H.; Hoch, J. A.; Hwa, T. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 67−72.
(11) Mora, T.; Walczak, A. M.; Bialek, W.; Callan, C. G. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 5405−5410.
(12) Dixit, P. D. *J. Chem. Phys.* **2013**, *138*, 184111.
(13) Dixit, P. D. *Phys. Chem. Chem. Phys.* **2015**, *17*, 13000−13005.
(14) Dixit, P. D. *Biophys. J.* **2013**, *104*, 2743−2750.
(15) Jaynes, E. T. *Annu. Rev. Phys. Chem.* **1980**, *31*, 579−601.
(16) Stock, G.; Ghosh, K.; Dill, K. A. *J. Chem. Phys.* **2008**, *128*, 194102.
(17) Otten, M.; Stock, G. *J. Chem. Phys.* **2010**, *133*, 034119.
(18) Dixit, P. D.; Dill, K. A. *J. Chem. Theory Comput.* **2014**, *10*, 3002−3005.
(19) Hazoglou, M. J.; Walther, V.; Dixit, P. D.; Dill, K. A. *J. Chem. Phys.* **2015**, *143*, 051104.
(20) Selimkhanov, J.; Taylor, B.; Yao, J.; Pilko, A.; Albeck, J.; Hoffmann, A.; Tsimring, L.; Wollman, R. *Science* **2014**, *346*, 1370−1373.
(21) Chodera, J. D.; Noé, F. *Curr. Opin. Struct. Biol.* **2014**, *25*, 135−144.
(22) Filyukov, A.; Karpov, V. Y. *J. Eng. Phys.* **1967**, *13*, 416−419.
(23) Cavagna, A.; Giardina, I.; Ginelli, F.; Mora, T.; Piovani, D.; Tavarone, R.; Walczak, A. M. *Phys. Rev. E* **2014**, *89*, 042707.
(24) McGibbon, R. T.; Pande, V. S. *J. Chem. Phys.* **2015**, *143*, 034109.
(25) Seifert, U. *Eur. Phys. J. B* **2008**, *64*, 423−431.
(26) Bicout, D.; Szabo, A. *J. Chem. Phys.* **1998**, *109*, 2325−2338.
(27) Hummer, G. *New J. Phys.* **2005**, *7*, 34.
(28) Glauber, R. J. *J. Math. Phys.* **1963**, *4*, 294.
(29) Mariz, A.; Herrmann, H.; de Arcangelis, L. *J. Stat. Phys.* **1990**, *59*, 1043−1050.
(30) Leffler, J. E.; Grunwald, E. *Rates and equilibria of organic reactions: As treated by statistical, thermodynamic and extrathermodynamic methods*; Courier Corp.: Mineola, NY, USA, 2013.

DOI: 10.1021/acs.jctc.5b00537
J. Chem. Theory Comput. 2015, 11, 5464−5472

(31) Michaelides, A.; Liu, Z.-P.; Zhang, C.; Alavi, A.; King, D. A.; Hu, P. *J. Am. Chem. Soc.* **2003**, *125*, 3704−3705.
(32) Marcus, R. A. *J. Phys. Chem.* **1968**, *72*, 891−899.
(33) Matouschek, A.; Fersht, A. R. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 7814−7818.
(34) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141−151.
(35) Hamelberg, D.; Mongan, J.; McCammon, J. A. *J. Chem. Phys.* **2004**, *120*, 11919−11929.
(36) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(37) Laio, A.; Gervasio, F. L. *Rep. Prog. Phys.* **2008**, *71*, 126601.
(38) Altis, A.; Otten, M.; Nguyen, P. H.; Hegger, R.; Stock, G. *J. Chem. Phys.* **2008**, *128*, 245102.
(39) Jain, A.; Stock, G. *J. Chem. Theory Comput.* **2012**, *8*, 3810−3819.
(40) Caticha, A.; Preuss, R. *Phys. Rev. E* **2004**, *70*, 046127.
(41) Caticha, A. Entropic Inference: Some Pitfalls and Paradoxes We Can Avoid. *Invited paper presented at MaxEnt 2012, 32nd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, arXiv preprint arXiv:1212.6967; 2012.
(42) Paulsson, J. *Nature* **2004**, *427*, 415−418.
(43) Baldi, P.; Frasconi, P.; Smyth, P. *Modeling the Internet and the Web: Probabilistic methods and algorithms*; John Wiley & Sons: Chichester, U.K., 2003; p 277, DOI: 10.1002/047086799X.index.

DOI: 10.1021/acs.jctc.5b00537
J. Chem. Theory Comput. 2015, 11, 5464−5472