

Fluctuation relations between hierarchical kinetically equivalent networks with Arrhenius-type transitions and their roles in systems and structural biology

De-Ming Deng, Yi-Ta Lu, and Cheng-Hung Chang

Institute of Physics, National Chiao Tung University, Hsinchu 300, Taiwan

(Received 22 March 2017; published 2 June 2017)

The legality of using simple kinetic schemes to determine the stochastic properties of a complex system depends on whether the fluctuations generated from hierarchical equivalent schemes are consistent with one another. To analyze this consistency, we perform lumping processes on the stochastic differential equations and the generalized fluctuation-dissipation theorem and apply them to networks with the frequently encountered Arrhenius-type transition rates. The explicit Langevin force derived from those networks enables us to calculate the state fluctuations caused by the intrinsic and extrinsic noises on the free energy surface and deduce their relations between kinetically equivalent networks. In addition to its applicability to wide classes of network related systems, such as those in structural and systems biology, the result sheds light on the fluctuation relations for general physical variables in Keizer's canonical theory.

DOI: [10.1103/PhysRevE.95.062401](https://doi.org/10.1103/PhysRevE.95.062401)

I. INTRODUCTION

Two contraction theories aimed at exploring hierarchical kinetic schemes in biological systems have recently been rather active. In systems biology (SYB), the attempt to extract the key dynamics of complex networks has boosted the study of lumping analysis (LA) originating in the petroleum industry [1–6]. In structural biology (STB), the quest to identify metastable states of macromolecules has given birth to methods like geometric clustering [7,8], kinetic clustering [9,10], and automatic algorithms for the discovery of kinetically metastable states [11,12]. Although the system scales and the objects to contract (biochemical species versus molecular metastable states) in those two fields are far apart from each other, the mathematics and strategies adopted are similar. For instance, the transition matrix \mathbf{T} used in STB is related to the rate constant matrix \mathbf{M} of the rate equation [see (1) below] in SYB by $\mathbf{T}(\tau) = e^{\mathbf{M}\tau}$ with a short time span τ [13]. The kinetic clustering in STB may contract geometrically dissimilar conformational states into a merged state [10], just as the approximate lumping in SYB may unify physically unrelated states into a lumped one [14]. When the contraction is based on temporal difference, separating states by time scale in the lumping process in SYB [6] is analogous to clustering metastable states via the eigenvectors of \mathbf{T} in the Perron cluster analysis method in STB [15]. The theories and algorithms independently developed in these two fields have collected fruitful information for understanding the relations between hierarchical dynamics, however, solely on the level of averaged behavior. On the level of fluctuations, Keizer pointed out some relation [16], which nevertheless is restricted to hierarchical dynamics connected by invertible transformations. For noninvertible lumping processes, such as those used in the above biochemical problems, it lacked a systematic study until the recently developed stochastic lumping analysis (SLA) [17].

A low-dimensional network prevalently used for accounting for the experimental data of a complex system can usually be coarse grained from other higher-dimensional networks. If two networks are equally good at describing those data, they are often anticipated to be kinetically equivalent (KE)

to each other, which is the core condition for establishing the theory of LA. Under this condition, the mean dynamics of two networks are “indistinguishable,” in that they exhibit identical deterministic dynamics after being projected to a desired reduced space. The traditional LA has succeeded in revealing a variety of relations between the mean dynamics of KE networks, which, however, evoke further questions about the relations between their fluctuations. Issues related to these questions include the possibility of resolving hidden network structures by fluctuations, the criteria for consistent fluctuations during model selection, and whether and when hierarchical KE models give the same conclusion to a fluctuation related system property. One step to clarify these questions has been achieved by the SLA, which has unraveled the duality of indistinguishability in the mean dynamics and fluctuations of KE networks under intrinsic noises and derived various fluctuation relations between those networks for extrinsic noises [18]. These results are attributed to the introduction of the lumping concept from the rate equation in the traditional LA into the chemical master equation and the stochastic differential equation. When fluctuations are small, the noise effects from the latter two equations can be unified in a Langevin formalism with a hierarchical stochastic force. Despite the generality of that formalism, the challenge of obtaining concrete Langevin forces from real systems makes it difficult to see how this abstract theory works in practice.

In this study, we derived that force from the free energy fluctuations for networks whose states undergo an Arrhenius-type transition, which is fundamentally rooted in statistical physics and appears widely in nature [16]. With this explicit force form, we extracted several fluctuation relations between KE networks. To give a concrete illustration, these relations were elucidated on the example of the conformational change of proteins, such as biological receptors. Fluctuation issues are of special interest for this kind of protein because they are crucial for the reliability of biological information processing [19–23]. These fluctuations can arise from the small copy number of a protein on a cell membrane (intrinsic noises) or the environmental stochasticity originating in the electrolyte solution (extrinsic noises) [18]. Theoretically, we proved several equalities and inequalities for hierarchical fluctuations caused

by different types of noise within the framework of Langevin formalism and the generalized fluctuation-dissipation theorem (FDT) [16], and provided various biological examples for these relations. Numerically, we carried out simulations in the equilibrium and nonequilibrium regimes of fluctuation distribution to confirm the analytically derived relations and explore further relations beyond them. On the one hand, the results unveil how fluctuations of different levels of contracted KE systems are or should be related, which gives a theoretical basis for the consistency of using kinetic schemes. On the other hand, they serve as a prototype example for understanding the fluctuation relations between hierarchical dynamics of general physical variables in Keizer's canonical theory. The evolution of fluctuations in that theory is determined by the linearized rate equations associated with an Arrhenius-type transition between states, which has the same mathematical structure as those discussed in Sec. II C below.

This work is structured as follows: For the theoretical part, in Sec. II A we discuss the relations between the covariances of driving and response in a single network subject to a noise. In Sec. II B these relations are proved to be structurally the same as those for judging the distinguishability between two networks. In Sec. II C we derive the Langevin force for rate equations whose rate constants have an Arrhenius form. In Sec. II D the SLA is used to exploit the fluctuation relations for KE networks whose states undergo the above Arrhenius-type transition. In Sec. II E these relations and the effective strengths of Langevin forces are investigated for several basic kinds of stochastic driving, including coherent, incoherent, and anticoherent free energy variations. For the numerical part, in Sec. III simulations are performed to demonstrate the above theoretical predictions. At the end we give a discussion in Sec. IV and a conclusion in Sec. V.

II. THEORY

A. Fluctuations in stochastic differential equations

Let kinetic scheme A be an n -dimensional network in which the evolution of states is governed by the rate equation

$$\frac{d\mathbf{N}}{dt} = \mathbf{M}\mathbf{N} \quad \text{or} \quad \frac{dN_i}{dt} = \sum_{j=1}^n k_{ji}N_j - k_{ij}N_i. \quad (1)$$

Here N_i is the population of the i th state, k_{ij} denotes the rate constant from state i to j , $k_{ii} \equiv 0$ for all i , and \mathbf{M} represents an indecomposable rate constant matrix with the entries $M_{ij} \equiv k_{ji} - \sum_{k=1}^n k_{ik}\delta_{ij}$, where δ_{ij} is a Kronecker δ , which is 1 for $i = j$ and 0 elsewhere. The total population, $\sum_{i=1}^n N_i$, in (1) is conserved, and \mathbf{M} is a noninvertible matrix of rank $n - 1$. If the system is subject to noises, the sure variable \mathbf{N} will be replaced by a random variable $\hat{\mathbf{N}} = \mathbf{N} + \delta\mathbf{N}$ with the fluctuation $\delta\mathbf{N}$ around the mean dynamics \mathbf{N} , where $\hat{\mathbf{N}}$ obeys the stochastic differential equation

$$\frac{d\hat{\mathbf{N}}}{dt} = \mathbf{M}\hat{\mathbf{N}} + \mathbf{f}, \quad (2)$$

with a stochastic force \mathbf{f} . Suppose this force is a Gaussian white noise obeying

$$\langle \mathbf{f}(t) \rangle = \mathbf{0} \quad \text{and} \quad \langle \mathbf{f}(t')\mathbf{f}^T(t) \rangle = \mathbf{\Gamma}\delta(t - t'), \quad (3)$$

with a symmetric, positive semidefinite, and generally time-dependent covariance matrix $\mathbf{\Gamma}$. Then the conditional covariance of $\delta\mathbf{N}$ is [17]

$$\boldsymbol{\sigma} \equiv \langle \delta\mathbf{N}\delta\mathbf{N}^T \rangle = \int_0^t e^{\mathbf{M}\tau} \mathbf{\Gamma} (e^{\mathbf{M}\tau})^T d\tau, \quad (4)$$

where $(e^{\mathbf{M}\tau})^T$ denotes the transpose of $e^{\mathbf{M}\tau}$ and $\boldsymbol{\sigma}$ and $\mathbf{\Gamma}$ stand for $\boldsymbol{\sigma}(t)$ and $\mathbf{\Gamma}(t - \tau)$, respectively. For simplicity, the time argument is not explicitly given in most variables throughout this paper, unless it could cause confusion, such as the time difference τ between $\boldsymbol{\sigma}(t)$ and $\mathbf{\Gamma}(t - \tau)$ in the same equation (4). As $\mathbf{\Gamma}$ is symmetric, the integral in (4) gives a symmetric $\boldsymbol{\sigma}$. Since \mathbf{f} and subsequently $\delta\mathbf{N}$ are Gaussian random variables, their second moments, $\mathbf{\Gamma}$ and $\boldsymbol{\sigma}$, are proper candidates for characterizing the fluctuations of driving source and response, respectively.

Taking the time derivative of $\boldsymbol{\sigma}$ in (4), it yields

$$\frac{d\boldsymbol{\sigma}}{dt} = \mathbf{M}\boldsymbol{\sigma} + \boldsymbol{\sigma}\mathbf{M}^T + \mathbf{\Gamma}, \quad (5)$$

where $\boldsymbol{\sigma} \equiv \boldsymbol{\sigma}(t)$ and $\mathbf{\Gamma} \equiv \mathbf{\Gamma}(t)$ do not have a time difference as those in (4). If $d\boldsymbol{\sigma}/dt = \mathbf{0}$, (5) is reduced to

$$\mathbf{M}\boldsymbol{\sigma} + \boldsymbol{\sigma}\mathbf{M}^T = -\mathbf{\Gamma}. \quad (6)$$

If the $\mathbf{\Gamma}$ in (5) varies with time, (6) can be used to calculate the equilibrium covariance, to which the $\boldsymbol{\sigma}$ in (5) will converge when $\mathbf{\Gamma}$ is frozen at an instantaneous time point t . At different t , (6) has different $\mathbf{\Gamma}$ and $\boldsymbol{\sigma}$. That time-dependent $\boldsymbol{\sigma}$ derived from (6) is distinct from the evolution of $\boldsymbol{\sigma}$ calculated from (5). Their deviation will be small, when the $\mathbf{\Gamma}$ in (5) varies slowly with time in comparison with the relaxation caused by \mathbf{M} , so that the dynamics is close to an adiabatic process. The deviation is a consequence of the fact that the $\boldsymbol{\sigma}$ in (4) and (5) depends on the history of $\mathbf{\Gamma}$, while that in (6) is merely the outcome of an instantaneous $\mathbf{\Gamma}$. Specifically, if the $\mathbf{\Gamma}$ in (5) is time independent, (6) is known as the generalized FDT at the equilibrium state of (1) [16] or the Lyapunov equation in the control theory. A time-independent $\mathbf{\Gamma}$ is the prerequisite for the $\hat{\mathbf{N}}$ in (2), or $\delta\mathbf{N}$, to have an equilibrium distribution.

Whether $\mathbf{\Gamma}$ and $\boldsymbol{\sigma}$ can be inferred from each other depends on their temporal properties. According to (6), if $\mathbf{\Gamma}$ is time independent and \mathbf{M} were invertible, $\mathbf{\Gamma}$ and $\boldsymbol{\sigma}$ would determine each other (see Appendix A). Since the \mathbf{M} in (1) is noninvertible, the $\boldsymbol{\sigma}$ in (6) can uniquely decide $\mathbf{\Gamma}$, but not vice versa (see Appendix A). According to (4) and (5), given a time t , " $\mathbf{\Gamma}(t') = \mathbf{0}$ for all $t' \in [0, t]$ " implies " $\boldsymbol{\sigma}(t') = \mathbf{0}$ for all $t' \in [0, t]$ " and vice versa (see Appendix A). For simplicity, it will be briefly expressed as

$$\mathbf{\Gamma} = \mathbf{0} \Leftrightarrow \boldsymbol{\sigma} = \mathbf{0}. \quad (7)$$

Here and throughout this paper, the symbol " $\mathbf{0}$ " denotes a null matrix. Relation (7) means that if the stochastic force is absent, an ensemble of systems governed by (2) and starting with an identical initial state will have zero fluctuations, and vice versa. Notice that (7) has $t' \in [0, t]$ individually on the both sides of the implication " \Leftrightarrow ." It is not a pointwise relation, and thus cannot be understood as " $\mathbf{\Gamma}(t') = \mathbf{0} \Leftrightarrow \boldsymbol{\sigma}(t') = \mathbf{0}$ " at any given $t' \in [0, t]$, because an instantaneous $\boldsymbol{\sigma}$ depends on the history of $\mathbf{\Gamma}$.

If only the variance of fluctuations is of concern, one can consider a diagonal matrix \mathbf{V} which has the same diagonal entries as σ . Since the transformation (4) from Γ to σ is linear, there exists a simple ordering relation (Appendix A),

$$\Gamma \bowtie \mathbf{0} \Rightarrow \sigma \bowtie \mathbf{0} \Rightarrow \mathbf{V} \bowtie \mathbf{0}. \quad (8)$$

Therein, the symbol \bowtie can be one of the following relations: “ $>$ ” (positive definite), “ $<$ ” (negative definite), “ \geq ” (positive semidefinite), “ \leq ” (negative semidefinite), or “ $=$ ” (null matrix). As in (7), each \bowtie in (8) holds individually for all $t' \in [0, t]$. Notably, the Γ in (7) and (8) are not restricted to the positive definite condition in (4), because they will be used to derive other similar relations in (18), (19), and (40) later.

B. Stochastic lumping analysis on stochastic differential equations

Let kinetic scheme A' be another network, which is n' -dimensional ($n' \leq n$) and whose rate equation and stochastic differential equation are

$$\frac{d\mathbf{N}'}{dt} = \mathbf{M}'\mathbf{N}' \quad \text{or} \quad \frac{dN'_a}{dt} = \sum_{b=1}^{n'} k'_{ba} N'_b - k'_{ab} N'_a, \quad (9)$$

$$\frac{d\hat{\mathbf{N}}'}{dt} = \mathbf{M}'\hat{\mathbf{N}}' + \mathbf{f}', \quad (10)$$

respectively, with a sure variable \mathbf{N}' , a random variable $\hat{\mathbf{N}}' = \mathbf{N}' + \delta\mathbf{N}'$, and an indecomposable rate constant matrix \mathbf{M}' . Here, $\delta\mathbf{N}'$ is the fluctuation of $\hat{\mathbf{N}}'$ around \mathbf{N}' and \mathbf{f}' is again a Gaussian white noise with the covariance Γ' , which generates a conditional covariance σ' , as the Γ in (3) generates the σ in (4).

If there exists an $n' \times n$ lumping matrix \mathbf{U} which can merge states in A into states in A' , such that \mathbf{M} and \mathbf{M}' obey

$$\mathbf{U}\mathbf{M} = \mathbf{M}'\mathbf{U}, \quad (11)$$

we say A is (exactly) lumpable and can be lumped into A' (Appendix B) [1, 17]. Under condition (11), A and A' are KE to each other, in that the contracted dynamics, $\mathbf{U}\mathbf{N}$, of (1) is identical with the dynamics, \mathbf{N}' , of the contracted system (9). With this property, the dynamics \mathbf{N} and \mathbf{N}' are called “indistinguishable.” In terms of the rate constants in \mathbf{M} and \mathbf{M}' , the KE condition (11) can alternatively be expressed as (B1) in Appendix B. For the problems considered below, it is sufficient to focus our discussion on the proper lumping, for which each column of \mathbf{U} is a standard unit vector [see example (B2) in Appendix B].

The KE condition (11) only gives a constraint between the mean dynamics (1) and (9), but not between the stochastic forces \mathbf{f} in (2) and \mathbf{f}' in (10). Thus, how this condition affects the fluctuations of $\mathbf{U}\hat{\mathbf{N}}$ and $\hat{\mathbf{N}}'$ is unclear from the traditional LA. The difference between these fluctuations can be quantified by [17]

$$\begin{aligned} \sigma_{\text{diff}} &\equiv \langle \delta\mathbf{N}'\delta\mathbf{N}'^T \rangle - \langle \delta(\mathbf{U}\mathbf{N})\delta(\mathbf{U}\mathbf{N})^T \rangle \\ &= \sigma' - \mathbf{U}\sigma\mathbf{U}^T = \int_0^t e^{\mathbf{M}'\tau} \Gamma_{\text{diff}}(e^{\mathbf{M}'\tau})^T d\tau, \end{aligned} \quad (12)$$

where σ_{diff} and Γ_{diff} denote $\sigma_{\text{diff}}(t)$ and $\Gamma_{\text{diff}}(t - \tau)$, respectively, with a time difference τ as that between the σ and Γ in

(4). Therein, the generally time-dependent symmetric matrices

$$\Gamma_{\text{diff}} \equiv \Gamma' - \mathbf{U}\Gamma\mathbf{U}^T \quad (13)$$

and σ_{diff} stand for the differences of driving and response, respectively, between two KE systems. The simple integral form in (12) is valid only for comparing two KE networks, because condition (11) has been used in the last equality of (12) [17].

Applying \mathbf{U} and \mathbf{U}^T to the left and right hand sides of (5) and making use of (11) yields the lumped fluctuation equation of A ,

$$\frac{d\mathbf{U}\sigma\mathbf{U}^T}{dt} = \mathbf{M}'\mathbf{U}\sigma\mathbf{U}^T + \mathbf{U}\sigma\mathbf{U}^T\mathbf{M}'^T + \mathbf{U}\Gamma\mathbf{U}^T. \quad (14)$$

In analogy to (5) of A , the σ' of A' follows

$$\frac{d\sigma'}{dt} = \mathbf{M}'\sigma' + \sigma'\mathbf{M}'^T + \Gamma'. \quad (15)$$

Subtracting (14) from (15), one obtains

$$\frac{d\sigma_{\text{diff}}}{dt} = \mathbf{M}'\sigma_{\text{diff}} + \sigma_{\text{diff}}\mathbf{M}'^T + \Gamma_{\text{diff}}, \quad (16)$$

which, under the condition $d\sigma_{\text{diff}}/dt = \mathbf{0}$, is reduced to

$$\mathbf{M}'\sigma_{\text{diff}} + \sigma_{\text{diff}}\mathbf{M}'^T = -\Gamma_{\text{diff}}. \quad (17)$$

Although \mathbf{M} is not explicitly seen in (12), (16), and (17), this matrix does affect those equations, because some parts of them can be reexpressed as \mathbf{M} by using (11).

Equations (12), (16), and (17), for comparing the fluctuations between two systems, have exactly the same structures as (4)–(6), for describing the fluctuations of a single system. Hence, one has

$$\Gamma_{\text{diff}} = \mathbf{0} \Leftrightarrow \sigma_{\text{diff}} = \mathbf{0}, \quad (18)$$

in analogy to (7), where $\Gamma_{\text{diff}} = \mathbf{0}$ and $\sigma_{\text{diff}} = \mathbf{0}$ hold individually for all $t' \in [0, t]$. Despite that analogy, the meanings of (18) and (7) are rather different. The equality $\Gamma = \mathbf{0}$ in (7) represents the absence of a stochastic force in a single system. By contrast, the equality $\Gamma_{\text{diff}} = \mathbf{0}$ in (18) refers to the indistinguishability between the covariances Γ and Γ' of two systems, whose stochastic forces are generally nonzero and can even be time dependent. To obtain equilibrium fluctuations, Γ and Γ' , rather than Γ_{diff} , must be constant in time. Let \mathbf{V}_{diff} be a diagonal matrix whose diagonal entries are the same as those of σ_{diff} . With the same argument as that for (8), one has

$$\Gamma_{\text{diff}} \bowtie \mathbf{0} \Rightarrow \sigma_{\text{diff}} \bowtie \mathbf{0} \Rightarrow \mathbf{V}_{\text{diff}} \bowtie \mathbf{0}, \quad (19)$$

where each \bowtie holds individually for all $t' \in [0, t]$.

As in general contraction theories, a lumping process will cause information loss, because \mathbf{U} is not an invertible transformation. Given a low-dimensional \mathbf{M}' , there are infinitely many \mathbf{U} and \mathbf{M} which satisfy (11). Even when \mathbf{U} has been selected, there still exist infinitely many \mathbf{M} which can be lumped into that \mathbf{M}' by the same \mathbf{U} . That is because condition (11) does not give a constraint to all rate constants in \mathbf{M} , as easily seen from its equivalent expression (B1). Likewise, for a given Γ' , infinitely many \mathbf{U} and Γ can fulfill $\Gamma_{\text{diff}} = \mathbf{0}$. Even when both Γ' and \mathbf{U} are known, Γ cannot be uniquely pinned down from $\Gamma_{\text{diff}} = \mathbf{0}$.

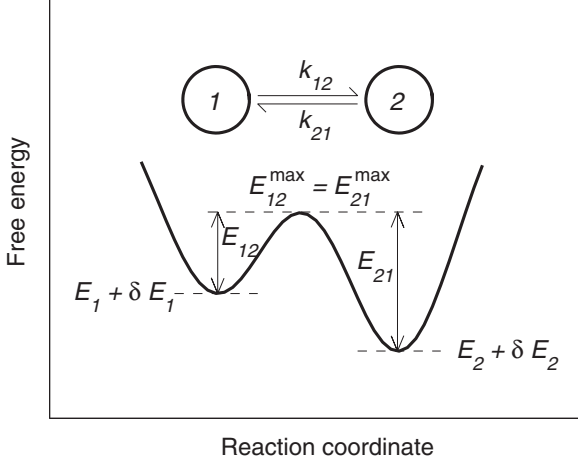


FIG. 1. A schematic free energy landscape has two minima, which may represent two conformations of a protein, such as the bound and the unbound state of a biological receptor or the open and the closed state of an ion channel. The transition dynamics between these two states can be described by a two-dimensional kinetic scheme, whose rate constants k_{ij} are related to the activation energies E_{ij} by the Arrhenius equation (21).

The theory discussed so far has focused on closed systems (1), which have a conserved total population. However, the formalism also applies to open systems which have some time-independent N_i in (1) and thus violate that conservation law. For open systems, the equilibrium states in the above derivation should be replaced by steady states. If Γ and Γ' vary with time, (2) and (10) are time-dependent Ornstein-Uhlenbeck stochastic processes, which are not stationary. If they are constant in time, these two equations describe the commonly known Ornstein-Uhlenbeck processes. In the following, we consider an Ornstein-Uhlenbeck process in a closed network system.

C. Langevin formalism for networks with an Arrhenius-type transition rate

A protein is often regarded as an n -state system, when its free energy surface can effectively be envisioned as having n minima (see an example for $n = 2$ in Fig. 1). Let E_i be the free energy of the i th state and suppose the transition between different states obey a rate equation as (1). Here, it is convenient to interpret the N_i in (1) as the probability of finding the protein in its i th state. If the free energies at these states vary stochastically, the energy vector $\mathbf{E} = [E_1, E_2, \dots, E_n]^T$ will fluctuate around its mean by $\delta\mathbf{E} = [\delta E_1, \delta E_2, \dots, \delta E_n]^T$. It leads to the fluctuations of the rate constants between different states, which has a similar effect as the fluctuations in potential barriers [24–26] and dynamical disorder [27,28], caused either by thermal noises or general stochastic processes. The energy fluctuations $\delta\mathbf{E}$ will then lead to the state fluctuations, $\delta\mathbf{N} = [\delta N_1, \delta N_2, \dots, \delta N_n]^T$, around the ensemble mean, $\mathbf{N} = [N_1, N_2, \dots, N_n]^T$.

To find a connection between $\delta\mathbf{E}$ and $\delta\mathbf{N}$, let us replace the mean values N_i and k_{ij} in (1) by $\hat{N}_i = N_i + \delta N_i$ and $\hat{k}_{ij} = k_{ij} + \delta k_{ij}$, respectively. In the first-order perturbation around the equilibrium states N_i^e of N_i , the values of δN_i

satisfy (Appendix C)

$$\frac{d\delta N_i}{dt} = \sum_{j=1}^n [(\delta k_{ji} N_j^e - \delta k_{ij} N_i^e) + (k_{ji} \delta N_j - k_{ij} \delta N_i)], \quad (20)$$

and are constrained by the conservation law $\sum_{i=1}^n \delta N_i = 0$. Assume that the rate constants in (20) have the Arrhenius form,

$$k_{ij} = A_{ij} e^{-\beta E_{ij}} = A_{ij} e^{-\beta(E_{ij}^{\max} - E_i)}, \quad (21)$$

where E_{ij} is the activation energy from state i to j , E_{ij}^{\max} denotes the elevation of the free energy barrier between these two states (Fig. 1), A_{ij} represents the pre-exponential factor, and $\beta = 1/(k_B T)$ stands for the inverse temperature, with the Boltzmann constant k_B and the absolute temperature T . Therein, E_{ij}^{\max} and A_{ij} are symmetric, where the former arises directly from its definition and the latter is a consequence of the microscopic reversibility [16]. Replacing E_{ij}^{\max} by $E_{ij}^{\max} + \delta E_{ij}^{\max}$ and E_i by $E_i + \delta E_i$ in (21), one obtains up to the first-order perturbation the equality

$$\begin{aligned} \delta k_{ij} &= -A_{ij} e^{-\beta(E_{ij}^{\max} - E_i)} \beta (\delta E_{ij}^{\max} - \delta E_i) \\ &= k_{ij} \beta (\delta E_i - \delta E_{ij}^{\max}). \end{aligned} \quad (22)$$

Additionally, the equilibrium probabilities N_i^e and N_j^e in (20) fulfill the detailed balance condition,

$$k_{ij} N_i^e - k_{ji} N_j^e = 0. \quad (23)$$

Substituting (22) and (23) into (20), it yields a Langevin-type equation (Appendix C),

$$\frac{d\delta\mathbf{N}}{dt} = \mathbf{M}(\delta\mathbf{N} + \beta \mathbf{D}_N^e \delta\mathbf{E}), \quad (24)$$

where \mathbf{D}_N^e is a diagonal matrix whose i th diagonal term is N_i^e . This equation describes the fluctuation of $\hat{\mathbf{N}}$ of a stochastic system around the mean \mathbf{N} of an ensemble of identical systems, when \mathbf{N} has reached equilibrium. Notice that even when \mathbf{N} is at equilibrium, the distribution of $\hat{\mathbf{N}}$ in the ensemble might not have arrived at an equilibrium shape. A commonly studied case is that all systems start with an identical initial state $\hat{\mathbf{N}}$ and accordingly have an initial distribution like a Dirac δ function. In such an ensemble, the covariance of $\delta\mathbf{N}$ will initially be conditioned at zero and evolve as (4). If these identical $\hat{\mathbf{N}}$ are exactly the equilibrium state of (1), the $\delta\mathbf{N}$ of each individual system in the ensemble is governed by (24).

Combining the stochastic term $\delta\mathbf{N}$ in (24) with the deterministic term \mathbf{N} in (1) leads to a stochastic differential equation as (2), with the specific form of stochastic force

$$\mathbf{f} = \beta \mathbf{M} \mathbf{D}_N^e \delta\mathbf{E}. \quad (25)$$

Although (25) looks complex, it is independent of $\delta\mathbf{N}$ and therefore is an additive noise, which does not suffer from the Ito-Stratonovich dilemma [18]. According to (25), different \mathbf{M} will generate different \mathbf{f} from the same $\delta\mathbf{E}$ and lead subsequently to different σ in (4). In this respect, measuring fluctuations $\delta\mathbf{N}$ induced by $\delta\mathbf{E}$ is instructive for understanding \mathbf{M} . This kind of relation arises naturally from the first-order expansion in general perturbation theories. It is reminiscent of the traditional FDT, which connects the microscopic fluctuation

tuations with the macroscopic relaxations (see the Discussion section).

In (24) and (25), $\delta\mathbf{N}$, $\delta\mathbf{E}$, and \mathbf{f} are n -dimensional vectors. Owing to the conservation of total probability, the degree of freedom of $\delta\mathbf{N}$ in (24) is reduced to $n - 1$. If the n components in $\delta\mathbf{E}$ are independent from one another, the degree of freedom of \mathbf{f} will also be $n - 1$, because \mathbf{f} is related to $\delta\mathbf{E}$ by the rank $n - 1$ matrix \mathbf{M} in (25). For $n = 2$, this reduction in the number of driving components is equivalent to taking the free energy difference $F_{12} \equiv E_1 - E_2$ as the single stochastic driving source in a two-state receptor system (Appendix C) [29]. Notably, the term δE_{ij}^{\max} in (22) has disappeared in (24). That means, to the first-order perturbation, the fluctuations of $\hat{\mathbf{N}}$ do not depend on the vibrations at the energy maxima along the transition paths, but only on those at the local minima, δE_j . Therefore, the state fluctuations driven by δE_{ij}^{\max} and δE_i are to the first-order perturbation identical with those driven solely by δE_i . This identity is rather general and independent of the type of noise assigned to $\delta\mathbf{E}$.

If $\delta\mathbf{E}$ is a Gaussian white noise with

$$\langle \delta\mathbf{E}(t') \rangle = \mathbf{0} \quad \text{and} \quad \langle \delta\mathbf{E}(t) \delta\mathbf{E}(t')^T \rangle = \hat{\mathbf{\Gamma}} \delta(t - t'), \quad (26)$$

with a symmetric $\hat{\mathbf{\Gamma}}$, the stochastic force \mathbf{f} in (25) will also be Gaussian and white and have the symmetric covariance

$$\bar{\mathbf{\Gamma}} \equiv \beta^2 \mathbf{M} \mathbf{D}_N^e \hat{\mathbf{\Gamma}} (\mathbf{M} \mathbf{D}_N^e)^T. \quad (27)$$

If $\hat{\mathbf{\Gamma}}$ and $\bar{\mathbf{\Gamma}}$ represent the strengths of an intrinsic noise, they are uniquely determined by the rate constants k_{ij} and the time-dependent N_i in (1) [16]. When N_i are at equilibrium, these covariances are constant in time, for which the distribution of $\delta\mathbf{N}$ will arrive at an equilibrium shape at $t \rightarrow \infty$. If $\hat{\mathbf{\Gamma}}$ and $\bar{\mathbf{\Gamma}}$ stand for the strengths of an extrinsic noise, they are generally time dependent and do not necessarily approach any constant matrices. Then the corresponding distribution of $\delta\mathbf{N}$ will never converge to an equilibrium shape, even though the mean \mathbf{N} of $\hat{\mathbf{N}}$ must always be at equilibrium, as assumed to obtain (24) and (25). More details about extrinsic and intrinsic noises are referred to in the Discussion section.

To explore the physical limits of biochemical signaling in chemotaxis, Bialek and Setayeshgar took into account the fluctuation effect from protein kinetics [29] in addition to that from ligand diffusion originally considered by Berg and Percell [19]. In that refined model, a biological receptor is assumed to be a two-state (bound and unbound) system, with k_+ (k_-) the rate constant from the unbound (bound) to the bound (unbound) state. Let \bar{n} be the equilibrium fractional occupancy of the receptor by the ligand. Then the thermal noise induced fluctuation $\delta\tilde{n}$ around \bar{n} will obey a Langevin equation [29]

$$\frac{d\delta\tilde{n}}{dt} = -(k_+c + k_-) \delta\tilde{n} + c(1 - \bar{n}) \delta k_+ - \bar{n} \delta k_-, \quad (28)$$

where δk_{\pm} are small changes in k_{\pm} and c is the ligand concentration. This scalar equation turns out to be a special case of the vector equation (24) with $n = 2$ (Appendix C).

Simple kinetic schemes are prevalently used to explain the fluctuation related physical properties of a complex system, such as the measurement uncertainty of chemotaxis explained by the two-state protein model in (28). If hierarchical KE

schemes predict inconsistent uncertainties, one would be puzzled by the question in which uncertainty we should believe. Therefore, coincident physical properties concluded from different KE schemes are crucial for whether we can rely on the results obtained from a certain scheme. If the physical property is defined on σ' and σ , that coincidence will then depend on whether the KE condition (11) will lead to the consistency in fluctuations, $\sigma_{\text{diff}} = \mathbf{0}$. This question will be clarified in the protein system in Sec. II C by the SLA theory in Sec. II B.

D. Fluctuations relations between hierarchical networks with an Arrhenius-type transition

Suppose the protein described by the above n -dimensional kinetic scheme, termed system A , can also be regarded as an n' -dimensional ($n' \leq n$) system A' , governed by a stochastic differential equation as (10). Following an analogous derivation as that for (25) and (27), one obtains the stochastic force \mathbf{f}' of A' and its covariance $\bar{\mathbf{\Gamma}}'$, given by

$$\mathbf{f}' = \beta \mathbf{M}' \mathbf{D}_{N'}^e \delta\mathbf{E}', \quad (29)$$

$$\bar{\mathbf{\Gamma}}' = \beta^2 \mathbf{M}' \mathbf{D}_{N'}^e \hat{\mathbf{\Gamma}}' (\mathbf{M}' \mathbf{D}_{N'}^e)^T. \quad (30)$$

Here, $\delta\mathbf{E}'$ is a Gaussian white noise with a covariance $\hat{\mathbf{\Gamma}}'$. That covariance generates a σ' , which is the σ in (4), with \mathbf{M} and $\mathbf{\Gamma}$ there replaced by \mathbf{M}' and $\bar{\mathbf{\Gamma}}'$, respectively. The distinguishability between σ' and the σ generated by the $\bar{\mathbf{\Gamma}}$ in (27) can be decided by $\sigma_{\text{diff}} \equiv \sigma' - \mathbf{U}\sigma\mathbf{U}^T$. If A and A' can equally well describe the experimentally observed mean behavior of the protein dynamics, they fulfill the KE condition (11). Under this condition, σ_{diff} can be simplified as

$$\sigma_{\text{diff}} = \int_0^t e^{\mathbf{M}'\tau} \bar{\mathbf{\Gamma}}_{\text{diff}} (e^{\mathbf{M}'\tau})^T d\tau, \quad (31)$$

with $\bar{\mathbf{\Gamma}}_{\text{diff}} \equiv \bar{\mathbf{\Gamma}}' - \mathbf{U}\bar{\mathbf{\Gamma}}\mathbf{U}^T$, as that in (12). The time derivative of (31) is

$$\frac{d\sigma_{\text{diff}}}{dt} = \mathbf{M}'\sigma_{\text{diff}} + \sigma_{\text{diff}}\mathbf{M}'^T + \bar{\mathbf{\Gamma}}_{\text{diff}}, \quad (32)$$

from which follows the relations (18) and (19), with $\mathbf{\Gamma}_{\text{diff}}$ there replaced by $\bar{\mathbf{\Gamma}}_{\text{diff}}$. As remarked in the text following (27), for general noises, $\bar{\mathbf{\Gamma}}$ and $\bar{\mathbf{\Gamma}}'$ are not necessarily time independent, although they are associated with an equilibrium system discussed here. For intrinsic noises, $\bar{\mathbf{\Gamma}}$ and $\bar{\mathbf{\Gamma}}'$ are time independent only when the system has arrived at equilibrium. For those constant $\bar{\mathbf{\Gamma}}$ and $\bar{\mathbf{\Gamma}}'$, (18) and (19) with $[0, t]$ extended to $[0, \infty)$ allow us to inspect the distinguishability, σ_{diff} , for equilibrium fluctuations.

In terms of the $\hat{\mathbf{\Gamma}}$ of $\delta\mathbf{E}$ in (27) and the $\hat{\mathbf{\Gamma}}'$ of $\delta\mathbf{E}'$ in (30), (31) becomes

$$\sigma_{\text{diff}} = \beta^2 \int_0^t e^{\mathbf{M}'\tau} \mathbf{M}' \hat{\mathbf{\Gamma}}_{\text{diff}} \mathbf{M}'^T (e^{\mathbf{M}'\tau})^T d\tau, \quad (33)$$

where the symmetric matrix

$$\hat{\mathbf{\Gamma}}_{\text{diff}} \equiv \mathbf{D}_{N'}^e \hat{\mathbf{\Gamma}}' \mathbf{D}_{N'}^e - \mathbf{U} \mathbf{D}_N^e \hat{\mathbf{\Gamma}} \mathbf{D}_N^e \mathbf{U}^T \quad (34)$$

is related to the $\bar{\mathbf{\Gamma}}_{\text{diff}}$ in (31) by

$$\bar{\mathbf{\Gamma}}_{\text{diff}} = \beta^2 \mathbf{M}' \hat{\mathbf{\Gamma}}_{\text{diff}} \mathbf{M}'^T. \quad (35)$$

In analogy to the σ_{diff} and Γ_{diff} in (12), the σ_{diff} and $\bar{\Gamma}_{\text{diff}}$ in (31) as well as the σ_{diff} and $\hat{\Gamma}_{\text{diff}}$ in (33) stand for matrices at times t and $t - \tau$, respectively. From (35) follows (Appendix D)

$$\hat{\Gamma}_{\text{diff}} \succcurlyeq \mathbf{0} \Rightarrow \bar{\Gamma}_{\text{diff}} \succcurlyeq \mathbf{0}, \quad (36)$$

where \succcurlyeq denotes “ \geq ,” “ \leq ,” or “ $=$,” which is valid pointwise at any instantaneous time.

The time derivative of σ_{diff} in (33) gives

$$\frac{d\sigma_{\text{diff}}}{dt} = \mathbf{M}'\sigma_{\text{diff}} + \sigma_{\text{diff}}\mathbf{M}'^T + \beta^2\mathbf{M}'\hat{\Gamma}_{\text{diff}}\mathbf{M}'^T. \quad (37)$$

A comparison between (37) and (32) reconfirms (35). Both (37) and (33) lead to the conclusion

$$\hat{\Gamma}_{\text{diff}} = \mathbf{0} \quad \begin{array}{l} \Rightarrow \\ \not\Leftarrow \end{array} \quad \sigma_{\text{diff}} = \mathbf{0}, \quad (38)$$

where both equalities hold individually for all $t' \in [0, t]$, which includes $[0, \infty)$, as explained in the text after (32). Here, the reason for the forward implication “ \Rightarrow ” is the same as that for (18). The failure of the reverse implication “ \Leftarrow ” is ascribed to the noninvertible \mathbf{M}' embedded in (33). The relation in (38) means that to observe indistinguishable state fluctuations $\sigma_{\text{diff}} = \mathbf{0}$, the free energy fluctuations $\delta\mathbf{E}$ and $\delta\mathbf{E}'$ are not necessarily “distinguishable,” $\hat{\Gamma}_{\text{diff}} = \mathbf{0}$. A plausible example for $\hat{\Gamma}_{\text{diff}} = \mathbf{0}$ is when $\hat{\Gamma}$ and $\hat{\Gamma}'$ are related by

$$\hat{\Gamma} = \mathbf{U}^T \hat{\Gamma}' \mathbf{U}, \quad (39)$$

which can be readily verified by inserting (39) into (34) and making use of the identity $\mathbf{D}_{\mathbf{N}}^e = \mathbf{U}\mathbf{D}_{\mathbf{N}}^e\mathbf{U}^T$. The form of (39) looks similar to, but in fact is different from, $\bar{\Gamma}_{\text{diff}} = \mathbf{0}$ with the $\bar{\Gamma}_{\text{diff}}$ in (31) or $\Gamma_{\text{diff}} = \mathbf{0}$ with the Γ_{diff} in (13), because the ordering of \mathbf{U}^T and \mathbf{U} in (39) is opposite that in those two identities. Owing to this ordering difference, (39) is an equality between $n \times n$ matrices, whereas $\bar{\Gamma}_{\text{diff}} = \mathbf{0}$ and $\Gamma_{\text{diff}} = \mathbf{0}$ are equalities between $n' \times n'$ matrices.

Although the claim in (38) for $\delta\mathbf{E}$ is weaker than that in (18) for \mathbf{f} , other claims for the former,

$$\hat{\Gamma}_{\text{diff}} \succcurlyeq \mathbf{0} \Rightarrow \sigma_{\text{diff}} \succcurlyeq \mathbf{0} \Rightarrow \mathbf{V}_{\text{diff}} \succcurlyeq \mathbf{0}, \quad (40)$$

are similar to those for the latter in (19), where each \succcurlyeq holds individually for all $t' \in [0, t]$, which contains $[0, \infty)$ as in (38). The claims in (40) stem directly from (36) and (19), with Γ_{diff} in (19) substituted by the special case $\bar{\Gamma}_{\text{diff}}$. A slight difference between (40) and (19) is that the inequalities “ $>$ ” and “ $<$ ” in the latter are absent in the former because they can be derived only under certain conditions. To see this subtlety, let us reexpress the diagonal terms of \mathbf{V}_{diff} as (Appendix D)

$$[\mathbf{V}_{\text{diff}}]_{aa} = \beta^2 \sum_{b=1}^d \lambda_b \int_0^t |\mathbf{e}_a^T \mathbf{M}' e^{\mathbf{M}'\tau} \boldsymbol{\xi}_b|^2 d\tau, \quad (41)$$

where $[\mathbf{V}_{\text{diff}}]_{aa}$, λ_b , and $\boldsymbol{\xi}_b$ refer to $[\mathbf{V}_{\text{diff}}(t)]_{aa}$, $\lambda_b(t - \tau)$, and $\boldsymbol{\xi}_b(t - \tau)$, respectively. Here, d is the total number of the nonzero eigenvalues of $\hat{\Gamma}_{\text{diff}}$ and $\boldsymbol{\xi}_b$ denotes the normalized eigenvector of the b th nonzero eigenvalue λ_b of $\hat{\Gamma}_{\text{diff}}$. Therefore, $e^{\mathbf{M}'\tau} \boldsymbol{\xi}_b$ is the point at which a trajectory of (9) starting from $\boldsymbol{\xi}_b$ will arrive after time τ . The term inside the absolute value, $|\cdot|$, is the a th component of the vector field of that point. With this dynamic interpretation, it is transparent

to see that (Appendix D)

$$\hat{\Gamma}_{\text{diff}} \succcurlyeq \mathbf{0} \Rightarrow \mathbf{V}_{\text{diff}} \succcurlyeq \mathbf{0}, \quad (42)$$

where each \succcurlyeq holds individually for all $t' \in [0, t]$, which includes $[0, \infty)$, as in (40). The σ_{diff} in (40) is absent in (42) because it does not have a simple expression as that for \mathbf{V}_{diff} in (41) to show its \succcurlyeq relation with $\hat{\Gamma}_{\text{diff}}$. In applications, (40) and (42) are basic criteria for judging the effective strengthes of different stochastic forces (see Fig. 4).

As a comparison, (18) and (19) are properties of \mathbf{f} and \mathbf{f}' , while (38), (40), and (42) are those of $\delta\mathbf{E}$ and $\delta\mathbf{E}'$. They are two natural ways to introduce stochastic driving sources to this system. Notice that (7), (8), (18), (19), (38), (40), and (42) have been shown to be true in an interval $[0, t]$ on both sides of each implication. However, one can easily see that the second implications of (8), (19), and (40) are even pointwise correct at any instantaneous time, as that in (36).

Apart from its contribution in clarifying (42), expression (41) also suggests possible variations of stochastic forces during environmental changes. If the covariances of $\delta\mathbf{E}$ and $\delta\mathbf{E}'$ are additively shifted by

$$\hat{\Gamma} \rightarrow \hat{\Gamma} + c\mathbf{1}\mathbf{1}^T \quad \text{and} \quad \hat{\Gamma}' \rightarrow \hat{\Gamma}' + c'\mathbf{1}'\mathbf{1}'^T, \quad (43)$$

σ_{diff} is invariant (Appendix D), because $\mathbf{1}\mathbf{1}^T$ and $\mathbf{1}'\mathbf{1}'^T$ lie in the null space of the integral transformation (33). Here, c (c') is a real number and $\mathbf{1}$ ($\mathbf{1}'$) denotes an n -(n' -)dimensional column vector with all components equal to 1. The two mappings in (43) form a two-dimensional gauge transformation, which does not modify the mean energies E_i , but only the magnitudes of the fluctuations δE_i around these means. For large c and c' , $\hat{\Gamma} + c\mathbf{1}\mathbf{1}^T$ and $\hat{\Gamma}' + c'\mathbf{1}'\mathbf{1}'^T$ will behave like $c\mathbf{1}\mathbf{1}^T$ and $c'\mathbf{1}'\mathbf{1}'^T$ with two comparatively small deviations $\hat{\Gamma}$ and $\hat{\Gamma}'$, respectively. Nevertheless, irrespective of how small these deviations are, σ_{diff} is merely determined by them, rather than by $c\mathbf{1}\mathbf{1}^T$ and $c'\mathbf{1}'\mathbf{1}'^T$.

The gauge characteristic in (43) is not a common property for general transformations. For instance, a transformation consisting of

$$\bar{\Gamma} \rightarrow \bar{\Gamma} + c\mathbf{1}\mathbf{1}^T \quad \text{and} \quad \bar{\Gamma}' \rightarrow \bar{\Gamma}' + c'\mathbf{1}'\mathbf{1}'^T \quad (44)$$

will not make σ_{diff} invariant, because, unlike (33), there is no noninvertible \mathbf{M}' in the transformation (31) to bring the $\mathbf{1}'\mathbf{1}'^T$ in (44) into a null space. Another simple transformation is the multiplicative change of the covariances of $\delta\mathbf{E}$ and $\delta\mathbf{E}'$,

$$\hat{\Gamma} \rightarrow c\hat{\Gamma} \quad \text{and} \quad \hat{\Gamma}' \rightarrow c'\hat{\Gamma}', \quad (45)$$

which does not give an invariant σ_{diff} either, except in the trivial case $c = c' = 1$ (Appendix D). In a real macromolecule, changing an environmental condition, such as the $p\text{H}$ value of the system, may readily affect the statistical properties of the stochastic forces and their corresponding $\hat{\Gamma}$ and $\hat{\Gamma}'$. The consistency condition $\sigma_{\text{diff}} = \mathbf{0}$ in hierarchical KE models should be retained under an environmental change, which can be preserved by (43), but not by (45). This analysis offers an additional information for targeting appropriate $\hat{\Gamma}$ and $\hat{\Gamma}'$.

E. Some basic types of stochastic forces

To gain more insight, let us consider some basic types of $\hat{\Gamma}$, which may be used to account for the stochastic properties of real biological systems. While the $\hat{\Gamma}$ of an intrinsic noise is uniquely determined by rate constants, that of an extrinsic noise can be rather diverse [18]. That diversity may arise from the noises of nonuniform migration of ions onto an ion channel under physiological conditions [30] or from the artificially generated noises for studying the energy transduction in ion pumps [26]. Even though the $\hat{\Gamma}$ of a system may be large and complex, properly shuffling the components of $\hat{\mathbf{N}}$ in (2) can often bring more correlated states together to make \mathbf{M} and $\hat{\Gamma}$ more block diagonal. The states in the same block will usually be merged into a common state during model contraction.

If this coarse-graining procedure is a lumping process, it is a contraction from a kinetic scheme A to its KE scheme A' . Suppose the $\hat{\Gamma}'$ of A' is diagonal and the $\hat{\Gamma}$ of A is block diagonal, represented by a direct sum of submatrices,

$$\hat{\Gamma} = \sum_{a=1}^{n'} \sum_{i \in S_a} \sum_{j \in S_a} \hat{\Gamma}_{ij} \mathbf{e}_i \mathbf{e}_j^T, \quad (46)$$

$$\hat{\Gamma}' = \sum_{a=1}^{n'} \hat{\Gamma}'_{aa} \mathbf{e}'_a \mathbf{e}'_a{}^T, \quad (47)$$

where \mathbf{e}_i and \mathbf{e}'_a are standard unit vectors [see the example in (B3)]. For the three-dimensional lumping example in (B2) in Appendix B, (46) and (47) are

$$\hat{\Gamma} = \begin{pmatrix} \hat{\Gamma}_{11} & \hat{\Gamma}_{12} & 0 \\ \hat{\Gamma}_{21} & \hat{\Gamma}_{22} & 0 \\ 0 & 0 & \hat{\Gamma}_{33} \end{pmatrix} \text{ and } \hat{\Gamma}' = \begin{pmatrix} \hat{\Gamma}'_{11} & 0 \\ 0 & \hat{\Gamma}'_{22} \end{pmatrix}. \quad (48)$$

The diagonal $\hat{\Gamma}'$ refers to the uncorrelated energy fluctuations between two lumped states, $a = 1$ and 2 , whereas the block diagonal $\hat{\Gamma}$ indicates that the correlated energy fluctuations are limited to the internal states of a lumped state (Appendix B).

The off-diagonal entries of a covariance matrix are often less in magnitude than the diagonal entries. Among others, correlation matrices are a typical example. If the $\hat{\Gamma}$ in (46) has this property, it can be reexpressed as

$$\hat{\Gamma}_{ij} = \sum_{a=1}^{n'} \gamma_{ij} \hat{\Gamma}'_{aa} \delta_{i \in S_a} \delta_{j \in S_a}. \quad (49)$$

Therein, $\delta_{i \in S_a}$ is an indicator function [see (B4) in Appendix B] and γ_{ij} are the correlation coefficients, where $|\gamma_{ij}| = 1$ for $i = j$ and $|\gamma_{ij}| \leq 1$ for $i \neq j$. With this property, the $\hat{\Gamma}$ in (48) will become

$$\hat{\Gamma} = \begin{pmatrix} \hat{\Gamma}'_{11} & \gamma_{12} \hat{\Gamma}'_{11} & 0 \\ \gamma_{21} \hat{\Gamma}'_{11} & \hat{\Gamma}'_{11} & 0 \\ 0 & 0 & \hat{\Gamma}'_{22} \end{pmatrix}. \quad (50)$$

Under this general setup, one can show that (Appendix E 1)

$$\hat{\Gamma} \text{ in (49) and } \hat{\Gamma}' \text{ in (47)} \Rightarrow \hat{\Gamma}_{\text{diff}} \geq \mathbf{0}. \quad (51)$$

This relation covers several fundamental types of energy fluctuations, as shown in Fig. 2. Together with (40), it leads to the following ordering rules for the magnitudes of state fluctuations:

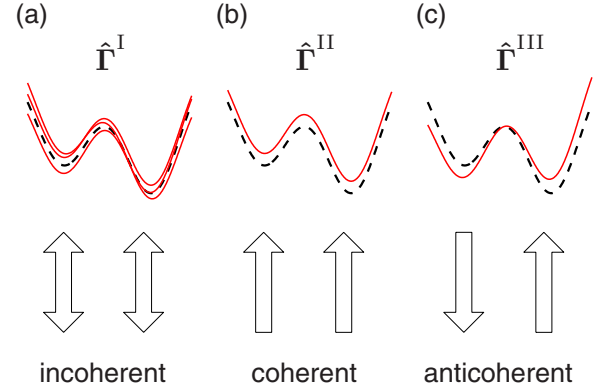


FIG. 2. During energy fluctuations, the energy landscape (black dashed curve) of the protein in Fig. 1 vibrates to different shapes (red solid curves). The two energy minima may fluctuate (a) independently described by an incoherent covariance $\hat{\Gamma}^{\text{I}}$, (b) in phase described by a coherent covariance $\hat{\Gamma}^{\text{II}}$, or (c) completely out of phase described by an anticoherent covariance $\hat{\Gamma}^{\text{III}}$.

(I) Incoherent driving. The simplest example of (51) is

$$\hat{\Gamma} = \mathbf{I}_{n \times n} \text{ and } \hat{\Gamma}' = \mathbf{I}_{n' \times n'} \Rightarrow \hat{\Gamma}_{\text{diff}} \geq \mathbf{0}, \quad (52)$$

where $\mathbf{I}_{n \times n}$ ($\mathbf{I}_{n' \times n'}$) is the n -(n')-dimensional identity matrix. In this case, all components in the $\delta \mathbf{E}$ of A ($\delta \mathbf{E}'$ of A') are uncorrelated, for which the example in (48) is reduced to

$$\hat{\Gamma} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \hat{\Gamma}' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (53)$$

This type of $\hat{\Gamma}$ ($\hat{\Gamma}'$) characterizes the incoherence between the driving sources, δE_i ($\delta E'_a$), at different energy states, such as those in Fig. 2(a). Since $\hat{\Gamma}_{\text{diff}} \geq \mathbf{0}$ in (52) implies $\mathbf{V}_{\text{diff}} \geq \mathbf{0}$ in (40), smaller KE networks will generate larger state fluctuations.

In reality, the diagonal forms in $\hat{\Gamma}$ and $\hat{\Gamma}'$ in (52) may occur when a protein is sensitive to some charged particles stochastically migrating onto it. If those particles approached the protein always from the same direction, δE_i and δE_j would be correlated and follow the linear or Marcus free energy relation [31]. But charged particles in electrolyte solutions usually arrive stochastically from arbitrary directions, which randomizes the correlation between δE_i and δE_j . Thus, the energy fluctuations at different states are likely incoherent and have diagonal $\hat{\Gamma}$ and $\hat{\Gamma}'$ as those in (52).

(II) Coherent driving. If $\gamma_{ij} = 1$ for all i and j at each a in (49), (51) is reduced to the relation (Appendix E 2)

$$\delta \mathbf{E} = \mathbf{U}^T \delta \mathbf{E}' \Rightarrow \hat{\Gamma}_{\text{diff}} = \mathbf{0}. \quad (54)$$

In this case, all internal states of a lumped state in A' have strong coherence in energy variations [Fig. 2(b)], so that the off-diagonal entries in each block of $\hat{\Gamma}$ are as large as the diagonal entries. With this property, the example in (48) becomes

$$\hat{\Gamma} = \begin{pmatrix} \hat{\Gamma}'_{11} & \hat{\Gamma}'_{11} & 0 \\ \hat{\Gamma}'_{11} & \hat{\Gamma}'_{11} & 0 \\ 0 & 0 & \hat{\Gamma}'_{22} \end{pmatrix} \text{ and } \hat{\Gamma}' = \begin{pmatrix} \hat{\Gamma}'_{11} & 0 \\ 0 & \hat{\Gamma}'_{22} \end{pmatrix}. \quad (55)$$

According to (54) and (40), A and A' will have indistinguishable covariances $\sigma_{\text{diff}} = \mathbf{0}$ and variances $\mathbf{V}_{\text{diff}} = \mathbf{0}$ in their state fluctuations.

Physically, the structure of $\hat{\Gamma}$ in (55) may occur when the internal states of a lumped state of a protein follow certain conformational symmetry. An example is the potassium ion channel with four identical gates, which have four configurations with one gate open and three gates closed [30]. These four configurations are indistinguishable in a patch-clamp recording. They belong to a global state [31] and each configuration is an internal state of that global state. If the four symmetric configurations respond coherently to an external stimulus, they form a 4×4 block in $\hat{\Gamma}$ with the same entry value in the block, similar to the example of the 2×2 block in (55).

(III) Anticoherent driving. If $\gamma_{ij} = 1$ for all $i = j$ and -1 for all $i \neq j$ at each a in (49), the property in (51) and $\mathbf{V}_{\text{diff}} \geq \mathbf{0}$ in (40) still hold. In this case, the example in (48) becomes

$$\hat{\Gamma} = \begin{pmatrix} \hat{\Gamma}'_{11} & -\hat{\Gamma}'_{11} & 0 \\ -\hat{\Gamma}'_{11} & \hat{\Gamma}'_{11} & 0 \\ 0 & 0 & \hat{\Gamma}'_{22} \end{pmatrix} \quad \text{and} \quad \hat{\Gamma}' = \begin{pmatrix} \hat{\Gamma}'_{11} & 0 \\ 0 & \hat{\Gamma}'_{22} \end{pmatrix}. \quad (56)$$

As the two variances, $\hat{\Gamma}'_{11}$ and $\hat{\Gamma}'_{22}$, are always non-negative, the diagonal (off-diagonal) entries in $\hat{\Gamma}$ and $\hat{\Gamma}'$ are positive or zero (negative or zero).

The negative entries in $\hat{\Gamma}$ represent the negative correlations between the fluctuating energies in the first two states, as those in Fig. 2(c). This may happen when two conformational states of a protein have opposite partial charges exposed to surroundings and respond oppositely to the local charge variation in an electrolyte solution.

(IV) Weakly coherent driving. A more realistic covariance $\hat{\Gamma}$ will have $0 < |\gamma_{ij}| < 1$ for $i \neq j$ in each block in (49), which lies between the above three limiting cases (I), (II), and (III). Since this setup also belongs to the category of (51), one has again $\mathbf{V}_{\text{diff}} \geq \mathbf{0}$ from (40).

The concrete examples discussed above reveal the effective strengths of the stochastic forces of several basic driving sources, which may readily be assumed to model a complex system. The general condition for γ_{ij} in (49) results in the inequality $\mathbf{\Gamma}_{\text{diff}} \geq \mathbf{0}$ in (51) and subsequently $\mathbf{V}_{\text{diff}} \geq \mathbf{0}$ from (40). It indicates the trend that a low-dimensional model will more easily overestimate the fluctuations of a real system, in comparison with a less coarse-grained model. This bias needs to be balanced by magnifying (diminishing) the values of all entries in $\mathbf{\Gamma}$ ($\mathbf{\Gamma}'$) to obtain consistent driving effects.

III. NUMERICAL SIMULATIONS

As applications, let us consider a four-dimensional network A , which can be lumped into a two-dimensional network A' by a lumping matrix \mathbf{U} , where the rate constants of A and A' as well as \mathbf{U} are given by

$$\begin{pmatrix} k_{11} & k_{12} & k_{13} & k_{14} \\ k_{21} & k_{22} & k_{23} & k_{24} \\ k_{31} & k_{32} & k_{33} & k_{34} \\ k_{41} & k_{42} & k_{43} & k_{44} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 5 & 15 \\ 0 & 0 & 5 & 15 \\ 10 & 20 & 0 & 0 \\ 10 & 20 & 0 & 0 \end{pmatrix},$$

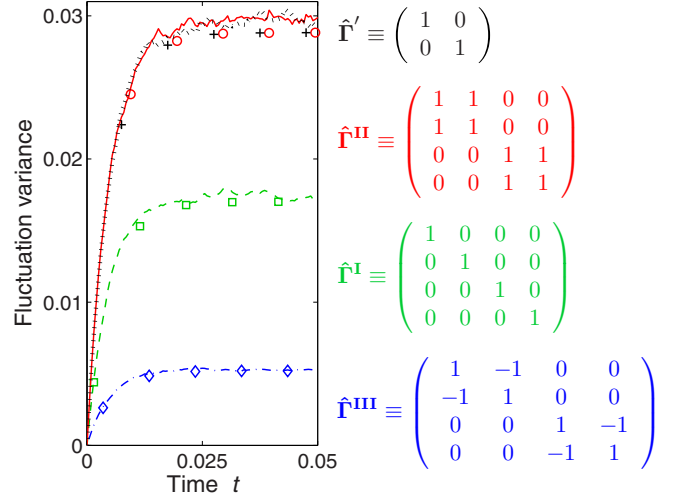


FIG. 3. The variance σ'_{11} of $\delta N'_1$ of a two-dimensional network subject to a stochastic force of covariance $\hat{\Gamma}'$ is compared with the lumped variance $[\mathbf{U}\sigma\mathbf{U}^T]_{11}$ of δN_1 and δN_2 of a four-dimensional KE network subject to a stochastic force of covariances $\hat{\Gamma}^I$, $\hat{\Gamma}^{II}$, and $\hat{\Gamma}^{III}$. Both the curves, simulated by the Langevin equation (24), and the symbols, theoretically evaluated by (4), follow the ordering rules predicted in (I)–(III).

$$\begin{pmatrix} k'_{11} & k'_{12} \\ k'_{21} & k'_{22} \end{pmatrix} = \begin{pmatrix} 0 & 20 \\ 30 & 0 \end{pmatrix}, \quad \text{and} \quad \mathbf{U} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}. \quad (57)$$

A and A' are subject to the stochastic forces in (25) and (29), respectively, with the covariances $\hat{\Gamma}'$ (black matrix) for the $\delta\mathbf{E}'$ in A' and $\hat{\Gamma}^I$, $\hat{\Gamma}^{II}$, and $\hat{\Gamma}^{III}$ (green, red, and blue matrices, respectively) for the $\delta\mathbf{E}$ in A , as given on the right-hand side of Fig. 3. The first diagonal entry of \mathbf{V}_{diff} , or σ_{diff} , is the difference between the σ'_{11} of A' and the $[\mathbf{U}\sigma\mathbf{U}^T]_{11}$ of A . Here, σ'_{11} is the variance of $\delta N'_1$ generated by $\hat{\Gamma}'$ and $[\mathbf{U}\sigma\mathbf{U}^T]_{11}$ is the sum of the second moments σ_{11} , σ_{12} , σ_{21} , and σ_{22} of δN_1 and δN_2 generated by $\hat{\Gamma}^I$, $\hat{\Gamma}^{II}$, or $\hat{\Gamma}^{III}$. On the left-hand side of Fig. 3, σ'_{11} is plotted in black (crosses and dotted curve), whereas $[\mathbf{U}\sigma\mathbf{U}^T]_{11}$ is depicted in green (squares and dashed curve), red (circles and solid curve), and blue (diamonds and dash-dotted curve). Therein, the symbols (crosses, circles, squares, and diamonds) are calculated from the integral expression (4), with $\mathbf{\Gamma}$ substituted by $\hat{\Gamma}^I$, $\hat{\Gamma}^{II}$, $\hat{\Gamma}^{III}$, and $\hat{\Gamma}'$. The curves (dotted, solid, dashed, and dash-dotted) are simulated by (24), which have been averaged over 10^4 realizations in a system of 10^2 proteins. Even though the curves averaged from this limited number of realizations still deviate slightly from the symbols, their ordering is already as clear as that of symbols and enables us to illustrate (40). Notice that during the time the variances of the fluctuation distributions in Fig. 3 approach their equilibrium values, the mean dynamics \mathbf{N} of the ensemble of systems stays always at equilibrium, because (24) is valid only for equilibrium systems. Owing to the conservation law $\sum_{i=1}^4 \delta N_i = \sum_{a=1}^2 \delta N'_a = 0$, σ'_{22} and $[\mathbf{U}\sigma\mathbf{U}^T]_{22}$ follow the same tracks as σ'_{11} and $[\mathbf{U}\sigma\mathbf{U}^T]_{11}$, respectively, and are therefore not depicted in the plot.

Figure 3 contains a lot of information about the fluctuation relations between different KE systems. First of all, both the curves and symbols confirm the orderings predicted by (40): For a comparison between $\hat{\Gamma} = \hat{\Gamma}^{\text{II}}$ and $\hat{\Gamma}'$, the coincidence between the red circles and black crosses indicates $\sigma_{\text{diff}} = \mathbf{0}$, as proved for the coherent driving in (II). For a comparison between $\hat{\Gamma} = \hat{\Gamma}^{\text{I}}$ ($\hat{\Gamma} = \hat{\Gamma}^{\text{III}}$) and $\hat{\Gamma}'$, the values of the black crosses are always larger than those of the green squares (blue diamonds), indicative of $\sigma_{\text{diff}} \geq \mathbf{0}$ and $\mathbf{V}_{\text{diff}} \geq \mathbf{0}$, as predicted from the incoherent driving in (I) [the antcoherent driving in (III)]. Besides confirming (I)–(III), Fig. 3 also unravels how different stochastic forces affect the state fluctuations in the same network. For example, the coherent driving by $\hat{\Gamma}^{\text{II}}$ is effectively stronger than the incoherent driving by $\hat{\Gamma}^{\text{I}}$ and than the antcoherent driving by $\hat{\Gamma}^{\text{III}}$, because the values of the red circles are larger than those of the green squares and those of the blue diamonds. This numerical result shows a simple tendency that more positive entries in $\hat{\Gamma}$ will give a stronger force, which proposes a hypothesis for further analytical justification. Moreover, the indistinguishable variances generated by $\hat{\Gamma} = \hat{\Gamma}^{\text{II}}$ (red circles) and $\hat{\Gamma}'$ (black crosses) are a good example for the claim that $\hat{\Gamma} = \mathbf{U}^T \hat{\Gamma}' \mathbf{U}$ in (39) will imply $\sigma_{\text{diff}} = \mathbf{0}$ in (38) discussed before. For simplicity of illustration, all $\hat{\Gamma}$ and $\hat{\Gamma}'$ used in the above simulations are time independent, which should not lead to the misunderstanding that (51) is valid only for constant $\hat{\Gamma}$ and $\hat{\Gamma}'$.

IV. DISCUSSIONS

Historically, the relation between σ , \mathbf{M} , and $\mathbf{\Gamma}$ has been intensively studied in statistical physics. Generally, $\mathbf{\Gamma}$ does not necessarily depend on \mathbf{M} , as when \mathbf{f} is an extrinsic noise. If $\mathbf{\Gamma}$ and \mathbf{M} originate from the same source, they will depend on each other [16], as when \mathbf{f} is an intrinsic noise. Irrespective of intrinsic or extrinsic, σ can be calculated by $\mathbf{\Gamma}$ and \mathbf{M} via (4) or (5). The FDT of Kubo is a theory for intrinsic noises, which relates the internal fluctuations of a system to the response of that system to a stimulus [32]. Traditional examples range over Nyquist’s theory, Brownian motion, and Onsager’s regression hypothesis. Certainly, Kubo’s FDT is stricter than the generalized FDT (6), because the former holds only for systems in which $\mathbf{\Gamma}$ depends on \mathbf{M} , while the latter is not restricted to this dependence. Keizer proposed a canonical theory for the fluctuations caused by elementary processes associated with an Arrhenius-type transition rate [16]. Because those fluctuations stem from the discrete nature of the microscopic constituents of a macroscopic system and $\mathbf{\Gamma}$ depends on \mathbf{M} , the canonical theory is aimed at intrinsic noises as well. It not only covers the equilibrium fluctuations discussed in Kubo’s FDT but also fluctuations far from equilibrium, including systems having time-dependent $\mathbf{\Gamma}$ and chaotic attractors. If the equilibrium fluctuations generated by (24) and (28) only need to obey the generalized FDT (6), these two Langevin equations can be used to discuss both intrinsic and extrinsic noises, as we do here. If they are additionally constrained by Kubo’s FDT, the analysis is restricted to intrinsic noises, as when (28) was derived in Ref. [29]. One might suspect whether the dependence of \mathbf{f} on

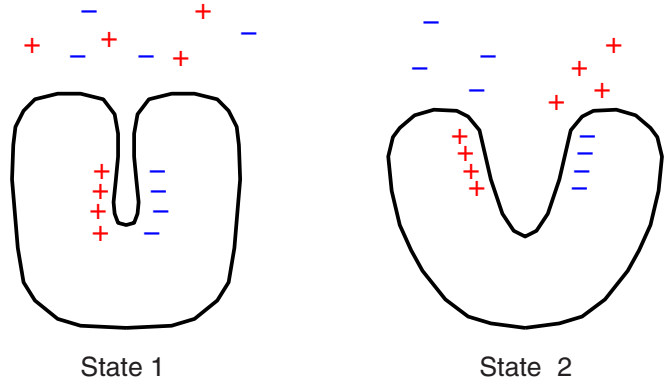


FIG. 4. A protein has two conformational states corresponding to the two energy minima in Fig. 1. As state 1 (state 2) here has less (more) partial charges exposing to the electrolyte solution, it is likely less (more) sensitive to the environmental noises. Thus, on average one has $|\delta E_1| < |\delta E_2|$. Under the same environmental noises, the ordering may readily change to $|\delta E_1| = |\delta E_2|$ or $|\delta E_1| > |\delta E_2|$, if the partial charge distribution or the macromolecular geometry is modified.

\mathbf{M} in (25) has restricted \mathbf{f} to be an intrinsic noise. It is not, because to be that noise, $\delta \mathbf{E}$ in \mathbf{f} must also depend on \mathbf{M} .

If a complex system subject to an intrinsic noise is modeled by two KE kinetic schemes, one can use the master equation to prove the indistinguishability between these two networks in all moments of the fluctuations of $\delta \mathbf{N}$, including the indistinguishable covariance, $\sigma_{\text{diff}} = \mathbf{0}$ [17]. If the noise is small, those fluctuations can also be calculated by stochastic differential equations, such as (2) and (10), whose $\mathbf{\Gamma}$ and $\mathbf{\Gamma}'$ are analytically known from \mathbf{M} and \mathbf{M}' , respectively. One can show that these $\mathbf{\Gamma}$ and $\mathbf{\Gamma}'$ will give $\mathbf{\Gamma}_{\text{diff}} = \mathbf{0}$ and, owing to (18), $\sigma_{\text{diff}} = \mathbf{0}$. It agrees with that concluded from the above master equation approach without the small noise assumption [17]. This indistinguishability in σ and σ' excludes the possibility of obtaining contradictory state fluctuations from different KE networks. It fortifies the legality of using simple kinetic schemes in the study of fluctuations generated by intrinsic noises.

By contrast, if \mathbf{f} in (2) represents an extrinsic noise [24,26,33], it summarizes the statistical properties of environmental disturbances and how the system responds to them. Even when the disturbance is the same, different systems may respond in different ways (Fig. 4). Since the environment can be rather complex and the corresponding $\mathbf{\Gamma}$ is decoupled from \mathbf{M} , it is much more difficult to derive $\mathbf{\Gamma}$ and $\mathbf{\Gamma}'$ to justify $\mathbf{\Gamma}_{\text{diff}} = \mathbf{0}$ than in the case of intrinsic noises. This difficulty is further complexified by the diverse meanings of “state” in a kinetic scheme. For instance, a state of a macromolecule usually corresponds to an effective minimum of a free energy landscape (Fig. 1), highly coarse grained from a much rugged potential energy landscape of atomic precision [34]. The coordinates of the free energy can be order parameters or progress variables, such as the numbers of native contacts, hydrogen bonds, and native dihedral angles [35]. The stochastic behavior of a state depends certainly on how the coordinates are selected. Thus, although $\mathbf{\Gamma}$ is to represent the stochastic driving effect outside the system, it also depends on

how the degrees of freedom inside the system is coarse grained. The combination of these external and internal complexities makes it nearly impossible to derive $\mathbf{\Gamma}$ for a real system to inspect whether $\mathbf{\Gamma}_{\text{diff}} = \mathbf{0}$, even for a short peptide or a small atom cluster [34]. However, despite that difficulty, $\mathbf{\Gamma}_{\text{diff}} = \mathbf{0}$ is a mathematical consequence of $\boldsymbol{\sigma}_{\text{diff}} = \mathbf{0}$ in (18). That is, $\mathbf{\Gamma}_{\text{diff}} = \mathbf{0}$ is true once two networks are KE to each other and have consistent covariances, $\boldsymbol{\sigma}_{\text{diff}} = \mathbf{0}$. But these two conditions are exactly what we usually request during model selection. Indeed, if \mathbf{f} represents an extrinsic noise, in practice we rarely consider it as a quantity to derive but mostly as a variable to tune to fit the observed experimental data. In this proposition, $\mathbf{\Gamma}_{\text{diff}} = \mathbf{0}$ is not a claim to be proved, but conversely serves as a criterion for identifying the desired \mathbf{M} and $\mathbf{\Gamma}$ from \mathbf{M}' and $\mathbf{\Gamma}'$.

As an example, suppose an ion channel of n states can be described by an n -dimensional kinetic scheme A . However, the binary switching of its conductance measured from a patch-clamp recording suggests it to be a two-state system. Those binary data can be used to construct the \mathbf{M}' , $\boldsymbol{\sigma}'$, and subsequently $\mathbf{\Gamma}'$ of a two-dimensional model A' (Appendix A). If both A and A' can correctly account for the dynamic features of those data, they are KE to each other and follow the KE condition (11). Mathematically, there exist infinitely many \mathbf{U} and \mathbf{M} which satisfy (11) for a given \mathbf{M}' , as noticed at the end of Sec. II B. This diversity will be largely reduced, if $\boldsymbol{\sigma}_{\text{diff}} = \mathbf{\Gamma}_{\text{diff}} = \mathbf{0}$, or equivalently $\boldsymbol{\sigma}' = \mathbf{U}\boldsymbol{\sigma}\mathbf{U}^T$ and $\mathbf{\Gamma}' = \mathbf{U}\mathbf{\Gamma}\mathbf{U}^T$, are considered, because \mathbf{M} is related to $\mathbf{\Gamma}$ and $\boldsymbol{\sigma}$ by (5) and (6). Noises are known to be useful for inferring with the hidden structures of networks, as those in detecting metastable states [36] and solving inverse problems [37,38]. Even in the above highly indefinite problem to infer A from A' , noise information, $\boldsymbol{\sigma}_{\text{diff}} = \mathbf{0}$, is still constructive.

Finally, there exists an intriguing correspondence between statistical physics and LA for networks whose states have an Arrhenius-type transition rate. Inserting the Arrhenius equation (21) into the rate constants in the KE condition (B1), it yields

$$e^{-\beta E'_{ab}} = \sum_{j \in S_b} (A_{ij}/A'_{ab}) e^{-\beta E_{ij}}. \quad (58)$$

This expression has the same form as the definition, $e^{-\beta F} = \sum \rho_i e^{-\beta \varepsilon_i}$, of a free energy F , with the system energies ε_i and their densities of state ρ_i . Just as the free energy is an effective energy of all system energies, the activation energy E'_{ab} from a lumped state a to another lumped state b can be interpreted as an effective energy of the activation energies E_{ij} from an internal state i of a to all internal states of b .

V. CONCLUSION

Motivated by its original goal to extract the key behaviors of complex networks, LA has become a theory for understanding the relations between the mean dynamics of KE networks. SLA takes into account stochasticity and generalizes these relations from the mean dynamics to the fluctuations of KE systems. In this work, we extended the SLA theory in a more systematical way and studied it on networks associated with the frequently encountered Arrhenius-type transition. The explicit form of the Langevin force derived from that

kind of transition allows us to deduce various fluctuation relations between hierarchical KE networks subject to intrinsic and various extrinsic noises. The result not only sets up a mathematical basis for justifying the consistency between the fluctuations among KE kinetic schemes, but also gives insight into the relations between the hierarchical fluctuations of general physical observables. For instance, the fluctuating δN under the fluctuating free energy in (24) is a special case of the fluctuating extensive variable under its fluctuating thermodynamically conjugate variable in Keizer's canonical theory [16]. As the canonical form of transition rate in that theory shares the same Arrhenius structure with (21), the fluctuation relations concluded here may be fundamental for general physical observables at different levels of description in nonequilibrium statistical physics.

ACKNOWLEDGMENTS

We thank G. Hu, J.-H. Lin, T. J. Kobayashi, J. Ohkubo, and L.-W. Yang for helpful discussions and the support of the National Center for Theoretical Sciences at Taiwan and of the Ministry of Science and Technology of Taiwan through Grant No. NSC 102-2112-M-009-012.

APPENDIX A

1. Let \mathbf{G} be an invertible matrix which can transform \mathbf{M} into a diagonal matrix $\mathbf{M}'' = \mathbf{G}\mathbf{M}\mathbf{G}^{-1}$. If \mathbf{G} and its transpose \mathbf{G}^T act on the both sides of (5), one has

$$\frac{d\boldsymbol{\sigma}''}{dt} = \mathbf{M}''\boldsymbol{\sigma}'' + \boldsymbol{\sigma}''\mathbf{M}''^T + \mathbf{\Gamma}'', \quad (A1)$$

with $\boldsymbol{\sigma}'' = \mathbf{G}\boldsymbol{\sigma}\mathbf{G}^T$ and $\mathbf{\Gamma}'' = \mathbf{G}\mathbf{\Gamma}\mathbf{G}^T$. Then the dynamics of the individual entries in this matrix equation can be decoupled and will follow the scalar equations

$$\frac{d\sigma''_{ij}}{dt} = \sigma''_{ij}(M''_{ii} + M''_{jj}) + \Gamma''_{ij}. \quad (A2)$$

If $\mathbf{\Gamma}$, and subsequently $\mathbf{\Gamma}''$, is time independent, the system will eventually reach its equilibrium state, whose $\boldsymbol{\sigma}''$ is determined by

$$\sigma''_{ij}(M''_{ii} + M''_{jj}) = -\Gamma''_{ij}. \quad (A3)$$

If \mathbf{M} and then \mathbf{M}'' were invertible, the $\boldsymbol{\sigma}''$ and $\mathbf{\Gamma}''$ in (A3) and subsequently $\boldsymbol{\sigma}$ and $\mathbf{\Gamma}$ would uniquely determine each other. Since \mathbf{M} in (1) is noninvertible and of rank $n - 1$, it has a zero eigenvalue. This will lead to a vanishing diagonal term, say M''_{kk} , in the diagonalized matrix \mathbf{M}'' . If $\boldsymbol{\sigma}'' = \mathbf{0}$, (A3) implies $\mathbf{\Gamma}'' = \mathbf{0}$, which is equivalent to the implication from $\boldsymbol{\sigma} = \mathbf{0}$ to $\mathbf{\Gamma} = \mathbf{0}$. Conversely, if $\mathbf{\Gamma}'' = \mathbf{0}$, all σ''_{ij} must be zero except for the indefinite σ''_{kk} , caused by $M''_{kk} = 0$. Although σ''_{kk} is the only unknown entry in $\boldsymbol{\sigma}''$, all entries in $\boldsymbol{\sigma} = \mathbf{G}^{-1}\boldsymbol{\sigma}''(\mathbf{G}^T)^{-1}$ will become unknown. Hence we have the relation

$$\mathbf{\Gamma}^* = \mathbf{0} \quad \not\Rightarrow \quad \boldsymbol{\sigma}^* = \mathbf{0} \quad (A4)$$

for the covariance $\boldsymbol{\sigma}^*$ of the equilibrium fluctuations of a system at a fixed $\mathbf{\Gamma}^*$. This pointwise relation derived from (6) is in contrast to the nonpointwise relation (7) obtained from (4) and (5).

2. If $\Gamma(t') = \mathbf{0}$ is valid for all $t' \in [0, t]$, $\sigma(t')$ in (5) does not change with time within that time interval, because $\sigma(t')$ is a conditional covariance starting with $\sigma(0) = \mathbf{0}$. Conversely, if $\sigma(t') = \mathbf{0}$ for all $t' \in [0, t]$, one obtains straightforwardly $\Gamma(t') = \mathbf{0}$ from (5) within the same time interval.

This result can also be intuitively understood from (4). The forward implication “ \Rightarrow ” in (7) is trivial from (4). For its reverse implication “ \Leftarrow ”, $\sigma(t') = \mathbf{0}$ at any $t' \in [0, t]$ implies a vanishing integrand in (4), which yields $\Gamma(t') = \mathbf{0}$ within that time interval. Here, $\sigma(t')$ is like the “area” bounded under the “curve” $H(\tau) \equiv e^{\mathbf{M}\tau} \Gamma(t' - \tau) (e^{\mathbf{M}\tau})^T$ in (4) and between the two ends at $\tau = 0$ and t' . If the area is zero for any interval $[0, t']$, $H(\tau)$ and subsequently $\Gamma(\tau)$ must vanish as well for all τ .

3. If \bowtie in (8) denotes “ $>$,” $\Gamma(t')$ is positive definite for all $t' \in [0, t]$ and therefore $\mathbf{v}\Gamma(t - \tau)\mathbf{v}^T > 0$ for any nonzero row vector \mathbf{v} and all $\tau \in [0, t]$. It implies that $\mathbf{v}e^{\mathbf{M}\tau}\Gamma(t - \tau)(e^{\mathbf{M}\tau})^T\mathbf{v}^T = \tilde{\mathbf{v}}\Gamma(t - \tau)\tilde{\mathbf{v}}^T > 0$ for any nonzero \mathbf{v} , because $\tilde{\mathbf{v}} \equiv \mathbf{v}e^{\mathbf{M}\tau}$ is again a nonzero vector, owing to the invertible matrix $e^{\mathbf{M}\tau}$. Therefore, $e^{\mathbf{M}\tau}\Gamma(t - \tau)(e^{\mathbf{M}\tau})^T$ is also a positive definite matrix. As $\sigma(t)$ in (4) is an integral of $e^{\mathbf{M}\tau}\Gamma(t - \tau)(e^{\mathbf{M}\tau})^T$ over the interval $[0, t]$, it is positive definite as well at time t . This is true also for $\sigma(t')$ at any time $t' \in [0, t]$, which can be obtained by shrinking the integration interval from $[0, t]$ to $[0, t']$. According to the property of positive definite, the diagonal elements of $\sigma(t')$ are positive, because $\mathbf{e}_i^T \sigma(t') \mathbf{e}_i > 0$ for all i , where \mathbf{e}_i are the standard unit vectors defined in the text before (B3). Consequently, the matrix $\mathbf{V}(t')$, which is the diagonal part of $\sigma(t')$, is also positive definite for all $t' \in [0, t]$.

If \bowtie in (8) represents “ $<$,” “ \geq ,” and “ \leq ,” the argument is the same. If \bowtie stands for “ $=$,” the claim that $\Gamma = \mathbf{0}$ implies $\sigma = \mathbf{0}$ is known from the forward part of (7).

APPENDIX B

If an $n' \times n$ lumping matrix \mathbf{U} is used to contract the n -dimensional \mathbf{N} in (1) of network A , the n states in A are first partitioned into n' sets S_a via the n' row vectors in \mathbf{U} , where $a = 1, \dots, n'$. Thereafter, all states in S_a are merged as the state a in A' and termed “the internal states” of a . In general, the n' -dimensional contracted dynamics $\mathbf{N}' = \mathbf{U}\mathbf{N}$ is not self-contained and, owing to a memory kernel, does not follow a simple linear equation like (9). If it does, we say A is lumpable and can be lumped into A' by \mathbf{U} , for which the \mathbf{M} of A , the \mathbf{M}' of A' , and \mathbf{U} obey the matrix condition (11). In terms of rate constants, this condition is equivalent to [17]

$$k'_{ab} = \sum_{j \in S_b} k_{ij} \quad (\text{B1})$$

for all $a, b \in \{1, 2, \dots, n'\}$ and all $i \in S_a$, with $k'_{aa} \equiv 0$.

As an example, let A be a system of three states $\{1, 2, 3\}$, which fulfills (1) with $n = 3$. If a lumping matrix

$$\mathbf{U} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{B2})$$

is used to contract A into A' , it will partition the states of A into two sets $S_1 = \{1, 2\}$ and $S_2 = \{3\}$. Then, state 1 and 2 of A will be merged into state 1' of A' , and state 3 of A will be

renamed as state 2' of A' . State 1 and 2 of A are the internal states of state 1' of A' . Since each column of (B2) is a standard unit vector, this matrix performs a proper lumping. After being contracted by \mathbf{U} , the dynamics $\mathbf{N}' = [N'_1, N'_2]^T$ of state 1' and 2' does not necessarily fulfill a simple equation like (9). If it does, A is lumpable.

Let \mathbf{e}_i (\mathbf{e}'_a) be an n -(n' -)dimensional standard unit column vector whose i th (a th) component is 1 and other components are 0. The product $\mathbf{e}_i \mathbf{e}'_a^T$ gives a matrix whose (i, j) th entry is 1 and other entries are 0. For instance, if we have $(i, j) = (3, 2)$ in a four-dimensional system, then

$$\mathbf{e}_3 \mathbf{e}'_2{}^T = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (\text{B3})$$

The k th column of \mathbf{U} is $\mathbf{U}\mathbf{e}_k$, which is some n' -dimensional standard unit vector \mathbf{e}'_a . The indicator function of a proper lumping matrix \mathbf{U} is defined as

$$\delta_{i \in S_a} \equiv \mathbf{e}'_a{}^T \mathbf{U}\mathbf{e}_i = \begin{cases} 1, & \text{if } i \in S_a, \\ 0, & \text{if } i \notin S_a, \end{cases} \quad (\text{B4})$$

which behaves like a Kronecker δ and indicates whether state i belongs to the set S_a .

APPENDIX C

1. Perturbing the rate equation (1) by replacing the rate constants k_{ij} by $k_{ij} + \delta k_{ij}$ and the probabilities N_i by $N_i + \delta N_i$, one obtains

$$\begin{aligned} & \frac{d(N_i + \delta N_i)}{dt} \\ &= \sum_{j=1}^n [(k_{ji} + \delta k_{ji})(N_j + \delta N_j) - (k_{ij} + \delta k_{ij})(N_i + \delta N_i)] \\ &= \sum_{j=1}^n [(k_{ji}N_j - k_{ij}N_i) + (\delta k_{ji}N_j - \delta k_{ij}N_i) \\ & \quad + (k_{ji}\delta N_j - k_{ij}\delta N_i) + (\delta k_{ji}\delta N_j - \delta k_{ij}\delta N_i)], \end{aligned} \quad (\text{C1})$$

with four parentheses in the last sum. The first one can be eliminated by (1) and the fourth one is the second-order term of the perturbation, which will be neglected because the fluctuations considered here are small. Thus it is left with (20).

2. Substituting (22) into (20) and considering the detailed balance condition (23), we obtain

$$\begin{aligned} \frac{d\delta N_i}{dt} &= \sum_{j=1}^n [N_j^e k_{ji} \beta (\delta E_j - \delta E_{ij}^{\max}) \\ & \quad - N_i^e k_{ij} \beta (\delta E_i - \delta E_{ij}^{\max}) + k_{ji} \delta N_j - k_{ij} \delta N_i] \\ &= \sum_{j=1}^n [\beta (k_{ji} N_j^e \delta E_j - k_{ij} N_i^e \delta E_i) \\ & \quad + k_{ji} \delta N_j - k_{ij} \delta N_i]. \end{aligned} \quad (\text{C2})$$

Rearranging the summands yields

$$\begin{aligned}
\frac{d\delta N_i}{dt} &= \sum_{j=1}^n [\beta k_{ji} N_j^c \delta E_j + k_{ji} \delta N_j] \\
&\quad - \sum_{j=1}^n k_{ij} \sum_{k=1}^n (\beta N_k^c \delta E_k + \delta N_k) \delta_{ik} \\
&= \sum_{j=1}^n [k_{ji} (\beta N_j^c \delta E_j + \delta N_j)] \\
&\quad - \sum_{k=1}^n k_{ik} \sum_{j=1}^n (\beta N_j^c \delta E_j + \delta N_j) \delta_{ik} \\
&= \sum_{j=1}^n \left[k_{ji} (\beta N_j^c \delta E_j + \delta N_j) \right. \\
&\quad \left. - \sum_{k=1}^n k_{ik} \delta_{ij} (\beta N_j^c \delta E_j + \delta N_j) \right]. \quad (C3)
\end{aligned}$$

Since $M_{ij} = k_{ji} - \sum_{k=1}^n k_{ik} \delta_{ij}$, as defined in (1), we arrive at the simple expression

$$\frac{d\delta N_i}{dt} = \sum_{j=1}^n [M_{ij} (\beta N_j^c \delta E_j + \delta N_j)], \quad (C4)$$

whose vector form is (24).

3. If a protein has two states 1 and 2, let us denote by k_+ (k_-) the rate constant from state 1 to 2 (2 to 1). When the protein is subject to noises, the fluctuations of the probabilities for finding states 1 and 2 are governed by a two-dimensional (24) with

$$\begin{aligned}
\mathbf{M} &= \begin{pmatrix} -k_- & k_+ \\ k_- & -k_+ \end{pmatrix}, \quad \mathbf{D}_N^c = \begin{pmatrix} N_1^c & 0 \\ 0 & N_2^c \end{pmatrix}, \\
\delta \mathbf{E} &= \begin{pmatrix} \frac{\delta k_-}{k_- \beta} + \delta E^{\max} \\ \frac{\delta k_+}{k_+ \beta} + \delta E^{\max} \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} -\delta k_- N_1^c + \delta k_+ N_2^c \\ \delta k_- N_1^c - \delta k_+ N_2^c \end{pmatrix}, \quad (C5)
\end{aligned}$$

where $\delta \mathbf{E}$ has been expressed as δk_{\pm} by (22). Since both components of \mathbf{f} are proportional to $\delta F_{12} = \delta E_1 - \delta E_2$, the two-dimensional stochastic source \mathbf{f} has the same degree of freedom as the one-dimensional δF_{12} .

For a receptor protein surrounded by ligands of concentration c mentioned in (28), let its bound state (unbound state) be the above state 1 (state 2). Then the k_+ and δk_+ in (C5) needs to be replaced by ck_+ and $c\delta k_+$. The N_1^c and δN_1 in the \mathbf{f} of (C5) are the equilibrium fractional occupancy \bar{n} and its fluctuation $\delta \bar{n}$ in (28). Inserting the specific matrices and vectors in (C5) into (24), the two components of (24) will become the same dynamics as (28) [29].

APPENDIX D

1. The reason for “ $\hat{\Gamma}_{\text{diff}} \succ \mathbf{0} \Rightarrow \beta^2 \mathbf{M}' \hat{\Gamma}_{\text{diff}} \mathbf{M}'^T \succ \mathbf{0}$ ” in (36) is similar to that for “ $\mathbf{\Gamma} \succ \mathbf{0} \Rightarrow e^{\mathbf{M}\tau} \mathbf{\Gamma} (e^{\mathbf{M}\tau})^T \succ \mathbf{0}$ ” in Appendix A. The main difference is that the noninvertible \mathbf{M}' in the former has replaced the invertible $e^{\mathbf{M}\tau}$ in the latter. As

a result, $\mathbf{v} \beta^2 \mathbf{M}' \hat{\Gamma}_{\text{diff}} \mathbf{M}'^T \mathbf{v}^T = \mathbf{v} \bar{\Gamma}_{\text{diff}} \mathbf{v}^T$ can be zero for some nonzero row vector \mathbf{v} , even when $\hat{\Gamma}_{\text{diff}} > \mathbf{0}$ or $< \mathbf{0}$. Thus, a positive definite (negative definite) $\hat{\Gamma}_{\text{diff}}$ in (35) does not necessarily imply a positive definite (negative definite) $\bar{\Gamma}_{\text{diff}}$. Consequently, “ $>$ ” and “ $<$ ” have to be excluded from \succ , which leads to \succsim in (36).

2. The \mathbf{V}_{diff} in (40) is the diagonal part of the σ_{diff} in (33), which can be decomposed as

$$\begin{aligned}
\mathbf{V}_{\text{diff}} &= \sum_{a=1}^{n'} \mathbf{e}'_a \mathbf{e}'_a{}^T \sigma_{\text{diff}} \mathbf{e}'_a \mathbf{e}'_a{}^T \\
&= \beta^2 \sum_{a=1}^{n'} \int_0^t \mathbf{e}'_a \mathbf{e}'_a{}^T e^{\mathbf{M}'\tau} \mathbf{M}' \hat{\Gamma}_{\text{diff}} \mathbf{M}'^T (e^{\mathbf{M}'\tau})^T \mathbf{e}'_a \mathbf{e}'_a{}^T d\tau, \quad (D1)
\end{aligned}$$

where $\mathbf{e}'_a \mathbf{e}'_a{}^T$ is as defined before (B3). Since $\hat{\Gamma}_{\text{diff}}$ is symmetric, it can be orthogonally diagonalized as

$$\hat{\Gamma}_{\text{diff}} = \mathbf{\Xi} \mathbf{D} \mathbf{\Xi}^T = \sum_{b=1}^d \mathbf{\Xi} (\mathbf{e}'_b \lambda_b \mathbf{e}'_b{}^T) \mathbf{\Xi}^T = \sum_{b=1}^d \xi_b \lambda_b \xi_b^T. \quad (D2)$$

Here, \mathbf{D} is an $n' \times n'$ diagonal matrix whose b th diagonal term is the b th eigenvalue λ_b of $\hat{\Gamma}_{\text{diff}}$, $\mathbf{\Xi}$ denotes an $n' \times n'$ orthogonal matrix whose b th column ξ_b is the b th normalized eigenvector of $\hat{\Gamma}_{\text{diff}}$, and d represents the number of nonzero λ_b with $0 \leq d \leq n'$, as defined in the main text. Substituting (D2) into (D1), the a th diagonal term of \mathbf{V}_{diff} becomes

$$\begin{aligned}
[\mathbf{V}_{\text{diff}}]_{aa} &= \beta^2 \int_0^t \sum_{b=1}^d \mathbf{e}'_a{}^T \mathbf{M}' e^{\mathbf{M}'\tau} \xi_b \lambda_b \xi_b^T (\mathbf{M}' e^{\mathbf{M}'\tau})^T \mathbf{e}'_a d\tau \\
&= \beta^2 \int_0^t \sum_{b=1}^d \lambda_b |\mathbf{e}'_a{}^T \mathbf{M}' e^{\mathbf{M}'\tau} \xi_b|^2 d\tau, \quad (D3)
\end{aligned}$$

where the commutation relation between $e^{\mathbf{M}'\tau}$ and \mathbf{M}' has been used in the last equality. Therein,

$$\begin{aligned}
\mathbf{e}'_a{}^T \mathbf{M}' e^{\mathbf{M}'\tau} \xi_b &= [\mathbf{M}' \mathbf{N}' |_{\mathbf{N}' = e^{\mathbf{M}'\tau} \xi_b}]_a \\
&= \left[\frac{d\mathbf{N}'}{d\tau} \Big|_{\mathbf{N}' = e^{\mathbf{M}'\tau} \xi_b} \right]_a \quad (D4)
\end{aligned}$$

has the dynamic interpretation in the text after (41). Exchanging the integral and the sum in (D3) leads to (41).

Expression (D3) can be used to reconfirm (40). To this end, recall that $\mathbf{V}_{\text{diff}} \equiv \mathbf{V}_{\text{diff}}(t)$, $[\mathbf{V}_{\text{diff}}]_{aa} \equiv [\mathbf{V}_{\text{diff}}]_{aa}(t)$, $\hat{\Gamma}_{\text{diff}} \equiv \hat{\Gamma}_{\text{diff}}(t - \tau)$, $\lambda_b \equiv \lambda_b(t - \tau)$, and $\xi_b \equiv \xi_b(t - \tau)$ in (D1), (D3), and (D4). If $\hat{\Gamma}_{\text{diff}}(t') \geq \mathbf{0}$ for all $t' \in [0, t]$, one has $\lambda_b \geq 0$ for all b in that time interval. Subsequently all summands in (D3) are non-negative at all $\tau \in [0, t]$, which yields $[\mathbf{V}_{\text{diff}}(t)]_{aa} \geq 0$ for all a . It also implies $[\mathbf{V}_{\text{diff}}(t')]_{aa} \geq 0$ for all a and all $t' \in [0, t]$, which can be obtained by simply shrinking the integration interval from $[0, t]$ to $[0, t']$. Therefore, it yields (40) for the case when \succsim denotes “ \geq .” For “ \leq ” and “ $=$,” the argument is the same.

3. The claim “ $\hat{\Gamma}_{\text{diff}} \succ \mathbf{0} \Rightarrow \mathbf{V}_{\text{diff}} \succ \mathbf{0}$,” with \succ denoting “ \geq ,” “ \leq ,” and “ $=$,” is clear from (40). For “ $\hat{\Gamma}_{\text{diff}}(t') > \mathbf{0}$,” one has $\lambda_b > 0$ for all b and all $t' \in [0, t]$. Then, (D3) implies $[\mathbf{V}_{\text{diff}}(t')]_{aa} > 0$ or $[\mathbf{V}_{\text{diff}}(t')]_{aa} = 0$ for all a at $t' = t$

and even at all $t' \in [0, t]$, as explained at the end of Appendix D. According to (D3), the sufficient and necessary condition for $[\mathbf{V}_{\text{diff}}]_{aa} = 0$ at a given a is that (D4) vanishes for all ξ_b . It is equivalent to $[dN'_a/d\tau]_{\mathbf{N}'=e^{M'\tau}\xi_b} = 0$ at all $\tau \in [0, t]$ and subsequently $[d\sum_{a' \neq a} N'_{a'}/d\tau]_{\mathbf{N}'=e^{M'\tau}\xi_b} = 0$ in the same time interval, due to the conservation of $\sum_{a'=1}^{n'} N'_{a'}$. That means the subpopulation in state a and that in the sum of all other states do not change with time τ during the evolution to equilibrium. Therefore, the ratio between initial subpopulations ($\sum_{a' \neq a} \Xi_{a'b}$)/ Ξ_{ab} must be identical with the ratio between equilibrium subpopulations ($\sum_{a' \neq a} N'_{a'}/N'_a \equiv \omega_a$, where the entry Ξ_{ab} of Ξ in (D2) is the a th component of ξ_b). This identity can be expressed as $\Omega_a \xi_b = 0$, with Ω_a a row vector whose a th component is $-\omega_a$ and other components are 1. Since the system has only one equilibrium state, the identity is valid for all ξ_b , which means $\Omega_a \Xi = \mathbf{0}$. As Ω_a is nonzero, $\Omega_a \Xi = \mathbf{0}$ holds only when Ξ is a noninvertible matrix, which would contradict its definition in (D2). Therefore, $[\mathbf{V}_{\text{diff}}]_{aa} = 0$ is excluded and only $[\mathbf{V}_{\text{diff}}]_{aa} > 0$ can follow from $\hat{\Gamma}_{\text{diff}} > \mathbf{0}$ in (D3). As a result, we have “ $\hat{\Gamma}_{\text{diff}} > \mathbf{0} \Rightarrow \mathbf{V}_{\text{diff}} > \mathbf{0}$.” The argument for \succ denoting “ $<$ ” is similar.

The KE systems excluded above are those which satisfy $[\mathbf{V}_{\text{diff}}(t)]_{aa} = 0$. According to (D3), they are systems which have (i) $d = 0$ or (ii) $d \geq 1$ and $|\cdot| = 0$ for all $b \in \{1, \dots, d\}$ and all $\tau \in [0, t]$. To fulfill (i), all eigenvalues of $\hat{\Gamma}_{\text{diff}}(t')$ must be 0, or equivalently $\hat{\Gamma}_{\text{diff}}(t') = \mathbf{0}$, for $t' \in [0, t]$. To satisfy (ii), (D4) implies that the trajectory of (9) starting from ξ_b should be time independent in its a th component, which includes the following three possibilities. First, if all components of that trajectory are time independent, ξ_b is proportional to the equilibrium state \mathbf{N}^e of (9), say $\xi_b = \rho \mathbf{N}^e$ with a normalized constant ρ , because $\mathbf{M}' e^{M'\tau} \xi_b = \rho \mathbf{M}' \mathbf{N}^e = \mathbf{0}$. In this case, we have not only $[\mathbf{V}_{\text{diff}}(t)]_{aa} = 0$ for a given a , but also $\mathbf{V}_{\text{diff}} = \mathbf{0}$. Since \mathbf{M}' can only have one equilibrium state, there exists merely one ξ_b , which cannot bring all summands in (D3) to 0, if $d \geq 2$. Therefore, in this case $[\mathbf{V}_{\text{diff}}]_{aa} = 0$ and $\mathbf{V}_{\text{diff}} = \mathbf{0}$ can occur only when $d \leq 1$. Second, if only some components, which contain the a th component, of the trajectory $e^{M'\tau} \xi_b$ in (D3) are time independent, the dynamics of these components may be disconnected from that of other components. This implies a decomposable matrix \mathbf{M}' and a nonergodic Markov chain generated by \mathbf{M}' [18], which, however, is of less interest and has been excluded in (1). Third, certain exceptional initial states of (9) may have only some time-independent components during the evolution, even when \mathbf{M}' is indecomposable. In this case, $[\mathbf{V}_{\text{diff}}]_{aa} = 0$ if a is one of those components. For instance, in a three-dimensional A' with $[k'_{12}, k'_{21}, k'_{13}, k'_{31}, k'_{23}, k'_{32}] = [2, 1, 1, 1, 0, 0]$, a trajectory \mathbf{N}' starting with the initial state $[N'_1(0), N'_2(0), N'_3(0)]^T = [1/4, \epsilon + 1/2, -\epsilon + 1/4]^T$ for $|\epsilon| < 1/4$ will fulfill $dN'/dt = [0, -\epsilon e^{-t}, \epsilon e^{-t}]^T$ for all t . Although the fluxes between different states in this system do not satisfy detailed balance before $t \rightarrow \infty$, the first component of \mathbf{N}' is always time independent, because the net flux from state 2 to 1 cancels with that from state 1 to 3. Interestingly, such local steady state occurs only when A' is lumpable again. In this example, A' can be lumped into a two-dimensional A'' with $[k''_{12}, k''_{21}] = [3, 1]$. Notice that this local steady state leads to $[\mathbf{V}_{\text{diff}}]_{aa} = 0$ for some a , but not $\mathbf{V}_{\text{diff}} = \mathbf{0}$.

4. Let us consider a special case,

$$c \mathbf{N}^e \mathbf{N}^{eT} \equiv \hat{\Gamma}_{\text{diff}}^{(0)}, \quad (\text{D5})$$

of (D2) with $d = 1$ and $\xi_1 = \rho \mathbf{N}^e$, where $c \equiv \lambda_1 \rho^2$ with λ_1 the eigenvalue of ξ_1 . The matrix $\hat{\Gamma}_{\text{diff}}^{(0)}$ lies in the null space of the linear transformation (33), because substituting it into the $\hat{\Gamma}_{\text{diff}}$ in (33) gives $\sigma_{\text{diff}} = \mathbf{0}$. If we shift $\hat{\Gamma}'$ by $c' \mathbf{1} \mathbf{1}^T$ and $\hat{\Gamma}$ by $c \mathbf{1} \mathbf{1}^T$ as those in (43), the $\hat{\Gamma}_{\text{diff}}$ in (34) will be converted into

$$\begin{aligned} & \mathbf{D}_{\mathbf{N}'}^e (\hat{\Gamma}' + c' \mathbf{1} \mathbf{1}^T) \mathbf{D}_{\mathbf{N}'}^e - \mathbf{U} \mathbf{D}_{\mathbf{N}}^e (\hat{\Gamma} + c \mathbf{1} \mathbf{1}^T) \mathbf{D}_{\mathbf{N}}^e \mathbf{U}^T \\ &= \hat{\Gamma}_{\text{diff}} + (c' - c) \hat{\Gamma}_{\text{diff}}^{(0)}. \end{aligned} \quad (\text{D6})$$

It will generate the same σ_{diff} as $\hat{\Gamma}_{\text{diff}}$ does, because $(c' - c) \hat{\Gamma}_{\text{diff}}^{(0)}$ also lies in the null space of (33).

If we change $\hat{\Gamma}'$ to $c' \hat{\Gamma}'$ and $\hat{\Gamma}$ to $c \hat{\Gamma}$ as those in (45), the $\hat{\Gamma}_{\text{diff}}$ in (34) will become

$$\mathbf{D}_{\mathbf{N}'}^e (c' \hat{\Gamma}') \mathbf{D}_{\mathbf{N}'}^e - \mathbf{U} \mathbf{D}_{\mathbf{N}}^e (c \hat{\Gamma}) \mathbf{D}_{\mathbf{N}}^e \mathbf{U}^T. \quad (\text{D7})$$

It will apparently lead to different σ_{diff} in (33) if c and c' are not 1.

APPENDIX E

1. For the special example of $\hat{\Gamma}$ and $\hat{\Gamma}'$ in (48), the decompositions in (46) and (47) mean

$$\begin{aligned} \hat{\Gamma} &= \sum_{i \in S_1} \sum_{j \in S_1} \hat{\Gamma}_{ij} \mathbf{e}_i \mathbf{e}_j^T + \sum_{i \in S_2} \sum_{j \in S_2} \hat{\Gamma}_{ij} \mathbf{e}_i \mathbf{e}_j^T \\ &= \begin{pmatrix} \hat{\Gamma}_{11} & \hat{\Gamma}_{12} & 0 \\ \hat{\Gamma}_{21} & \hat{\Gamma}_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \hat{\Gamma}_{33} \end{pmatrix}, \end{aligned} \quad (\text{E1})$$

$$\begin{aligned} \hat{\Gamma}' &= \sum_{a=1}^{n'} \hat{\Gamma}'_{aa} \mathbf{e}'_a \mathbf{e}'_a{}^T = \hat{\Gamma}'_{11} \mathbf{e}'_1 \mathbf{e}'_1{}^T + \hat{\Gamma}'_{22} \mathbf{e}'_2 \mathbf{e}'_2{}^T \\ &= \begin{pmatrix} \hat{\Gamma}'_{11} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \hat{\Gamma}'_{22} \end{pmatrix}. \end{aligned} \quad (\text{E2})$$

For general $\hat{\Gamma}$ in (46), the decomposition will be

$$\begin{aligned} [\mathbf{U} \mathbf{D}_{\mathbf{N}}^e \hat{\Gamma} \mathbf{D}_{\mathbf{N}}^e \mathbf{U}^T]_{ab} &= \mathbf{e}'_a{}^T \mathbf{U} \mathbf{D}_{\mathbf{N}}^e \hat{\Gamma} \mathbf{D}_{\mathbf{N}}^e \mathbf{U}^T \mathbf{e}'_b \\ &= \sum_{c=1}^{n'} \sum_{i \in S_c} \sum_{j \in S_c} \hat{\Gamma}_{ij} \mathbf{e}'_a{}^T \mathbf{U} \mathbf{D}_{\mathbf{N}}^e \mathbf{e}_i \mathbf{e}_j^T \mathbf{D}_{\mathbf{N}}^e \mathbf{U}^T \mathbf{e}'_b \\ &= \sum_{c=1}^{n'} \sum_{i \in S_c} \sum_{j \in S_c} \hat{\Gamma}_{ij} N_i^e N_j^e \mathbf{e}'_a{}^T \mathbf{U} \mathbf{e}_i \mathbf{e}_j^T \mathbf{U}^T \mathbf{e}'_b \\ &= \sum_{c=1}^{n'} \sum_{i \in S_c} \sum_{j \in S_c} \hat{\Gamma}_{ij} N_i^e N_j^e \delta_{i \in S_a} \delta_{j \in S_b} \\ &= \sum_{i \in S_a} \sum_{j \in S_b} N_i^e \hat{\Gamma}_{ij} N_j^e \delta'_{ab} \\ &= \sum_{i \in S_a} \sum_{j \in S_b} N_i^e N_j^e \gamma_{ij} \hat{\Gamma}'_{aa} \delta'_{ab}, \end{aligned} \quad (\text{E3})$$

where (49) has been used in the last equality. Therein, δ_{ab} denotes a Kronecker δ , which is 1 for $a = b$ and 0 elsewhere.

Likewise, for general $\hat{\Gamma}'$ in (47), the decomposition will give

$$\begin{aligned} [\mathbf{D}_{N'}^e \hat{\Gamma}' \mathbf{D}_{N'}^e]_{ab} &= \mathbf{e}_a^T \mathbf{D}_{N'}^e \hat{\Gamma}' \mathbf{D}_{N'}^e \mathbf{e}_b' = \sum_{c=1}^{n'} (N_c^e)^2 \hat{\Gamma}'_{cc} \mathbf{e}_c^T \mathbf{e}_c' \mathbf{e}_c^T \mathbf{e}_b' \\ &= \sum_{c=1}^{n'} (N_c^e)^2 \hat{\Gamma}'_{cc} \delta'_{ac} \delta'_{bc} = (N_a^e)^2 \hat{\Gamma}'_{aa} \delta'_{ab}. \end{aligned} \quad (\text{E4})$$

Substituting (E3) and (E4) into (34) and using the condition (49) again yields

$$\begin{aligned} [\hat{\Gamma}'_{\text{diff}}]_{ab} &= [\mathbf{D}_{N'}^e \hat{\Gamma}' \mathbf{D}_{N'}^e]_{ab} - [\mathbf{U} \mathbf{D}_N^e \hat{\Gamma}' \mathbf{D}_N^e \mathbf{U}^T]_{ab} \\ &= \left[(N_a^e)^2 - \sum_{i \in S_a} \sum_{j \in S_b} N_i^e N_j^e \gamma_{ij} \right] \hat{\Gamma}'_{aa} \delta'_{ab} \\ &\geq \left[(N_a^e)^2 - \sum_{i \in S_a} \sum_{j \in S_b} N_i^e N_j^e \right] \hat{\Gamma}'_{aa} \delta_{ab}. \end{aligned} \quad (\text{E5})$$

From the second line of (E5) we know $\hat{\Gamma}'_{\text{diff}}$ is diagonal. Owing to $\sum_{i \in S_a} N_i^e = N_a^e$, $\sum_{j \in S_b} N_j^e = N_b^e$, and the property of δ_{ab} , the term $[\cdot]$ in the last line of (E5) vanishes, which implies $[\hat{\Gamma}'_{\text{diff}}]_{ab} \geq 0$. Consequently, the diagonal matrix $\hat{\Gamma}'_{\text{diff}}$ has only non-negative entries and thus is positive semidefinite.

2. The claim in (54) is a special case of (51) with $\gamma_{ij} = 1$ for all i and j in (49), as (55) is a special example of (48). For those γ_{ij} , the second line of (E5) is equal to its third line, which is 0, as calculated above, and thus implies $\hat{\Gamma}'_{\text{diff}} = \mathbf{0}$. Alternatively, one can obtain (54) by inserting $\delta \mathbf{E} = \mathbf{U}^T \delta \mathbf{E}'$ there into (26),

$$\begin{aligned} \hat{\Gamma}' \delta(t' - t'') &= \langle \delta \mathbf{E}(t') \delta \mathbf{E}^T(t'') \rangle \\ &= \langle \mathbf{U}^T \delta \mathbf{E}'(t') \delta \mathbf{E}'^T(t'') \mathbf{U} \rangle = \mathbf{U}^T \hat{\Gamma}' \mathbf{U} \delta(t' - t''), \end{aligned} \quad (\text{E6})$$

because that implies $\hat{\Gamma}' - \mathbf{U}^T \hat{\Gamma}' \mathbf{U} = \mathbf{0}$ in (39) and subsequently $\hat{\Gamma}'_{\text{diff}} = \mathbf{0}$.

-
- [1] J. Wei and J. C. W. Kuo, *Ind. Eng. Chem. Fundam.* **8**, 114 (1969).
[2] G. Li and H. Rabitz, *Chem. Eng. Sci.* **44**, 1413 (1989).
[3] J. Tóth, G. Li, H. Rabitz, and A. S. Tomlin, *SIAM. J. Appl. Math.* **57**, 1531 (1997).
[4] M. S. Okino and M. L. Mavrouniotis, *Chem. Rev.* **98**, 391 (1998).
[5] A. N. Gorban, O. Radulescu, and A. Y. Zinovyev, *Chem. Eng. Sci.* **65**, 2310 (2010).
[6] J. C. Liao and E. N. Lightfoot Jr., *Biotechnol. Bioeng.* **31**, 869 (1988).
[7] M. E. Karpen, D. T. Tobias, and C. L. Brooks III, *Biochemistry* **32**, 412 (1993).
[8] B. de Groot, X. Daura, A. Mark, and H. Grubmüller, *J. Mol. Biol.* **309**, 299 (2001).
[9] V. Schultheis, T. Hirschberger, H. Carstens, and P. Tavan, *J. Chem. Theory Comput.* **1**, 515 (2005).
[10] F. Noé, I. Horenko, C. Schütte, and J. C. Smith, *J. Chem. Phys.* **126**, 155102 (2007).
[11] J. D. Chodera, N. Singhal, V. S. Pande, K. A. Dill, and W. C. Swope, *J. Chem. Phys.* **126**, 155101 (2007).
[12] *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*, edited by G. R. Bowman, V. S. Pande, and F. Noé (Springer, New York, 2014).
[13] F. Noé and S. Fischer, *Curr. Opin. Struct. Biol.* **18**, 154 (2008).
[14] J. C. W. Kuo and J. Wei, *Ind. Eng. Chem. Fundam.* **8**, 124 (1969).
[15] C. Schütte, A. Fischer, W. Huisinga, and P. Deuffhard, *J. Comput. Phys.* **151**, 146 (1999).
[16] J. Keizer, *Statistical Thermodynamics of Non-equilibrium Processes* (Springer-Verlag, New York, 1987).
[17] D. M. Deng and C. H. Chang, *J. Chem. Phys.* **142**, 184103 (2015).
[18] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 2007).
[19] H. C. Berg and E. M. Purcell, *Biophys. J.* **20**, 193 (1977).
[20] D. Fuller, W. Chen, M. Adler, A. Groisman, H. Levine, W.-J. Rappel, and W. F. Loomis, *Proc. Natl. Acad. Sci. USA* **107**, 9656 (2010).
[21] K. Kaizu, W. H. de Ronde, J. Paijmans, K. Takahashi, F. Tostevin, and P. R. ten Wolde, *Biophys. J.* **106**, 976 (2014).
[22] W. Bialek, *Biophysics: Searching for Principles* (Princeton University Press, Princeton, NJ, 2012).
[23] R. G. Endres, *Physical Principles in Sensing and Signaling: With an Introduction to Modeling in Biology* (Oxford University Press, Oxford, 2013).
[24] D. L. Stein, C. R. Doering, R. G. Palmer, J. L. van Hemmen, and R. M. McLaughlin, *J. Phys. A* **23**, L203 (1990).
[25] D. R. Glowacki, J. N. Harvey, and A. J. Mulholland, *Nat. Chem.* **4**, 169 (2012).
[26] H. V. Westerhoff, T. Y. Tsong, P. B. Chock, Y. D. Chen, and R. D. Astumian, *Proc. Natl. Acad. Sci. USA* **83**, 4734 (1986).
[27] R. Zwanzig, *Acc. Chem. Res.* **23**, 148 (1990).
[28] H. P. Lu, L. Xun, and X. S. Xie, *Science* **282**, 1877 (1998).
[29] W. Bialek and S. Setayeshgar, *Proc. Natl. Acad. Sci. USA* **102**, 10040 (2005).
[30] B. Hille, *Ion Channels of Excitable Membranes* (Sinauer Associates, Sunderland, 2001).
[31] M. B. Jackson, *Molecular and Cellular Biophysics* (Cambridge University Press, Cambridge, England, 2006).
[32] R. Kubo, *Rep. Prog. Phys.* **29**, 255 (1966).
[33] C.-H. Chang and T. Y. Tsong, *Phys. Rev. E* **69**, 021914 (2004).
[34] D. J. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, England, 2004).
[35] S. V. Krivov and M. Karplus, *Proc. Natl. Acad. Sci. USA* **101**, 14766 (2004).
[36] A. Alemany, A. Mossa, I. Junier, and F. Ritort, *Nat. Phys.* **8**, 688 (2012).
[37] J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, *Phys. Rev. Lett.* **104**, 058701 (2010).
[38] Z. Zhang, Z. Zheng, H. Niu, Y. Mi, S. Wu, and G. Hu, *Phys. Rev. E* **91**, 012814 (2015).