JOURNAL OF MATHEMATICAL PHYSICS **48**, 013303 (2007)

On detailed balance and reversibility of semi-Markov processes and single-molecule enzyme kinetics

Hongyun Wang[a)]
Department of Applied Mathematics and Statistics, University of California, Santa Cruz, California 95064

Hong Qian[b)]
Department of Applied Mathematics, University of Washington, Seattle, Washington 98195

(Received 17 July 2006; accepted 13 December 2006; published online 31 January 2007)

Semi-Markov processes have found increasing applications in modeling the kinetics of single enzyme molecules. Detailed balance is a widely accepted condition for Markov models of closed chemical systems and well known to be equivalent to the reversibility of a stationary Markov process. We show that for a semi-Markov process detailed balance is only a necessary condition, but not sufficient, for its time reversibility. The statistical independence between the transition direction and the sojourn time is also necessary. We show that the direction-time independence naturally arises from the exit problem of Markov models for enzyme kinetics with detailed balance. Detailed balance and the direction-time independence together are equivalent to the time reversibility of a stationary semi-Markov process. Applications of the present theory to single-molecule enzymology are also presented.
 [DOI: 10.1063/1.2432065]

I. INTRODUCTION

All molecular processes are stochastic in nature. The mathematical foundation of chemical reactions in condensed phases and aqueous solutions has been firmly established in the work of Kramers in terms of diffusion processes.[17,9] Discrete state, continuous time Markov processes have also found wide applications in chemical reaction dynamics and biochemical kinetics.[30,6,21]

There is a fundamental difference between chemical reactions in closed systems and in open systems.[22,10,26] The stationary states in the former are chemical equilibria with no net energy flow in any part of the system. They are time-reversible stochastic processes with detailed balance. This sets a stringent mathematical constraint on any realistic stochastic model for a closed chemical system. For diffusion processes defined by stochastic differential equations, this leads to the well-known fluctuation dissipation relations.[2,25] For discrete Markov processes this is reflected in the influential detailed balance condition.[31,18,30] In contrast, chemical and biochemical reactions in an open systems have continuous energy input and dissipation. The stationary states of such systems are known as nonequilibrium steady states (NESSs).[10,1,26] They are statistically time irreversible. Recent studies have shown that systems like such have distinguished positive *entropy productions* which quantitatively characterize the dissipation.[11]

In recent years, biophysical techniques in laboratories have made it possible to follow the stochastic dynamics of single biological macromolecules one at a time.[32,20] The data from these experiments demand a more systematic and stochastic approach to modeling and data interpretation. In particular, a new class of stochastic models in terms of semi-Markov processes (sMPs),

[a)]Electronic mail: hongwang@ams.ucsc.edu
[b)]Electronic mail: qian@amath.washington.edu

also known as Markov renewal processes (MRPs), has emerged.[24,27,14,29,15,28] sMPs are a very general class of stochastic models with well-known applications in operational research, queuing theory, and risk analysis.[13,19]

The natural mathematical question, therefore, is how to introduce the detailed balance and time reversibility to sMP. It turns out that detailed balance is no longer sufficient for the time reversibility of sMP. An additional condition we call the direction-time independence, motivated by enzyme kinetic studies, is also required. The direction-time independence is known in operation research literature.[3]

A finite sMP can be completely characterized by a semi-Markov kernel $Q(i,j,t), i \neq j$, defined as the probability that the first jump out of state i has occurred by time t and the destiny of the first jump is state j given that the system arrives at state i at time zero. $Q(i,j,\infty)$ is the probability that the destiny of the first jump out of state i is state j regardless of the waiting time. Associated with each sMP is an embedded Markov chain (eMC) with its transition probability $p_{ij} \equiv Q(i,j,\infty)$. In the present work, we establish several rigorous results on detailed balance and the time reversibility for sMP, as well as their relationship with statistical mechanics.

- We define detailed balance for a sMP as follows. If at the stationary state the net probability flux between any pair of states is zero, then we say that the sMP satisfies detailed balance.
- We show that detailed balance for a sMP is equivalent to detailed balance for its eMC.
- We show that a sufficient and necessary condition for a stationary sMP to be time reversible is detailed balance and $Q(i,j,t)=p_{ij}Q_i(t)$, which we shall call *direction-time independence*.
- We show that the existence of a potential function together with Boltzmann's relation is equivalent to the time reversibility of a sMP. The time reversibility is also equivalent to the symmetricity of certain linear operator associated with the probability evolution equation.
- We show that the direction-time independence is true in the exit problem associated with a class of reversible Markov processes motivated by enzyme kinetics.

We shall also demonstrate that detailed balance alone is not sufficient for the time reversibility of a stationary sMP. Notice that the direction-time independence is automatically satisfied in all Markov processes where detailed balance and time reversibility are known to be equivalent.

The paper is organized as follows. In Sec. II, we briefly summarize key known results on detailed balance and time reversibility for Markov chains. We shall also give a brief introduction to sMP. We then introduce the definition of detailed balance for sMP, and prove four theorems on detailed balance, time reversibility, Boltzmann's relation, and symmetry of probability evolution operator. In Sec. III, we provide further insight of the problem by studying a Markov model with detailed balance. The model is motivated by studies in single-molecule enzymology.[32,27,16] We show that the direction-time independence arises necessarily from systems like such, and prove another theorem.

II. SEMI-MARKOV PROCESSES: TIME REVERSIBILITY AND DETAILED BALANCE

A. Finite Markov chains: Time reversibility and detailed balance

Let $X=\{X_n; n \geqslant 0\}$ be an irreducible finite-state Markov chain with transition probability p_{ij}, $1 \leqslant i, j \leqslant N$. It can be shown that it has a unique stationary probability distribution π_i satisfying $\Sigma_{i=1}^N \pi_i p_{ij} = \pi_j$, where $\pi_i > 0$ and $\Sigma_{i=1}^N \pi_i = 1$.[7] A Markov chain is called detailed balanced if $\pi_i p_{ij} = \pi_j p_{ji}$. We then have the following theorem.

Theorem 1: *The following four statements are all equivalent to detailed balance.*

1. *The transition probabilities satisfy a loop law. That is, for any closed loop path* $\{i_0 i_1 \cdots i_n\}$ *where* $i_n = i_0$, *the relation below is satisfied,*

$$\frac{p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}}{p_{i_1 i_0} p_{i_2 i_1} \cdots p_{i_n i_{n-1}}} = 1. \tag{1}$$

2. *The stationary Markov chain with distribution* π_i *and transition probability* p_{ij} *is time reversible. Mathematically, the time reversibility is defined as*

$$\Pr\{(X_m X_{m+1} \cdots X_{m+n}) = (x_0 x_1 \cdots x_n)\} = \Pr\{(X_{m+n} X_{m+n-1} \cdots X_m) = (x_0 x_1 \cdots x_n)\} \tag{2}$$

for all $m \geqslant 0$, $n \geqslant 0$, *and* $(x_0, x_1, \ldots, x_n)$. *Basically, the time reversibility means that we shall see the same statistical behavior whether we read the time sequence* $\{X_n, n \geqslant 0\}$ *forward or backward.*

3. *There exists a potential function* G_i *such that for any pair of connected states* $G_i - G_j = \ln(p_{ij}/p_{ji})$ *and the stationary distribution* π_i *satisfies* $\pi_i/\pi_j = e^{-(G_i - G_j)}$.

4. *A quantity called entropy production is zero,*

$$\frac{1}{2}\sum_{i,j=1}^{N} (\pi_i p_{ij} - \pi_j p_{ji}) \ln\left(\frac{\pi_i p_{ij}}{\pi_j p_{ji}}\right) = 0. \tag{3}$$

See Ref. 11 for a proof of this theorem. The proof of the first three parts of this theorem can be found in many texts on reversible Markov chains, e.g., Ref. 12.

B. Markov renewal processes and semi-Markov processes

We introduce Markov renewal processes and sMPs according to Ref. 4.

Definition 1: *The stochastic process* $(X,T) = \{X_n, T_n; n \geqslant 0\}$ *is called a Markov renewal process provided that*

$$\Pr\{X_{n+1} = j, T_{n+1} - T_n \leqslant t | X_0, \ldots, X_n; T_0, \ldots, T_n\} = \Pr\{X_{n+1} = j, T_{n+1} - T_n \leqslant t | X_n\} \tag{4}$$

for all $n \geqslant 0$, $1 \leqslant j \leqslant N$, *and* $t \geqslant 0$, *where* N *is the number of possible states in the process. Here* n *is the index representing the number of jumps,* X_n *is the destiny of the nth jump, and* T_n *is the time at which the nth jump occurs. Basically, in a Markov renewal process, the stochastic evolution depends only on the current state and the time elapsed since arriving at the current state.* □

For $i \neq j$ and $t > 0$, let us introduce

$$Q(i,j,t) \equiv \Pr\{X_{n+1} = j, T_{n+1} - T_n \leqslant t | X_n = i\}, \tag{5}$$

which is called the semi-Markov kernel. We shall also denote

$$q_{ij}(t) \equiv \frac{\mathrm{d}}{\mathrm{d}t} Q(i,j,t), \quad p_{ij} \equiv Q(i,j,\infty), \quad Q_i(t) \equiv \sum_{l=1, l \neq i}^{N} Q(i,l,t), \tag{6}$$

$q_{ij}(t)\mathrm{d}t$ is the probability of jumping out of state i in time $[t, t+\mathrm{d}t)$ for the first time since arriving at state i at time zero and landing at destiny state j. p_{ij} is the probability of jumping from state i to state j regardless of the waiting time. $Q_i(t)$ is the cumulative distribution of the waiting time at state i regardless of the jumping destiny state. Note that $Q_i(t)$ depends only on i. We call the Markov chain with transition probability p_{ij} the eMC. Note that the eMC always satisfies $p_{ii} = 0$. A sMP is called irreducible if its eMC is irreducible.

Definition 2: *The stochastic jump process* $Y = \{Y_t; t \geqslant 0\}$ *defined as*

$$Y_t = X_n \quad \text{if } T_n \leqslant t < T_{n+1} \tag{7}$$

is called the sMP associated with the Markov renewal process (X,T). □

Definition 3: *The hazard function,*[23]

$$\beta_{ij}(t) \equiv \frac{q_{ij}(t)}{1 - Q_i(t)}, \tag{8}$$

is the conditional rate of jumping from state i *to state* j *at time* t *given that no jump occurs in* $[0,t)$.

The hazard function $\beta_{ij}(t)$ will be used in the evolution equation of age structure below. It is clear that

$$q_{ij}(t) = \beta_{ij}(t)\exp\left(-\int_0^t \sum_{l=1,l\neq i}^{N} \beta_{il}(s)\mathrm{d}s\right). \tag{9}$$

Notice that a sMP is Markov if and only if $\beta_{ij}(t)=\beta_{ij}$, independent of t.

C. Stationary solution and age structure

Consider an irreducible sMP. Let $r_i(\tau,t)$ be the time-dependent "age structure" of the sMP at state i, i.e., $r_i(\tau,t)\mathrm{d}\tau \equiv \Pr\{X_n=i, \tau<t-T_n\leqslant\tau+\mathrm{d}\tau\}$, where n is determined by $T_n\leqslant t<T_{n+1}$. The equation for the evolution of $r_i(\tau,t)$ is

$$\frac{\partial r_i(\tau,t)}{\partial t} = -\frac{\partial r_i(\tau,t)}{\partial \tau} - \left(\sum_{j=1,j\neq i}^{N} \beta_{ij}(\tau)\right) r_i(\tau,t), \tag{10}$$

$$r_i(0,t) = \int_0^\infty \left(\sum_{j=1,j\neq i}^{N} \beta_{ji}(s) r_j(s,t)\right)\mathrm{d}s. \tag{11}$$

Note that if all $\beta_{ij}(\tau)$'s are constant (i.e., it is a Markov process), then one can integrate the above equations over all age: $R_i(t)=\int_0^\infty r_i(\tau,t)\mathrm{d}\tau$ is the probability of being in state i at time t and $R_i(t)$'s satisfy the standard forward equation for a Markov process with transition rates β_{ij}.

The stationary age structure $r^*(\tau)$ can be solved from Eq. (10). Setting the left side of Eq. (10) to zero, multiplying by the integration factor, integrating with respect to τ, and using Eqs. (8) and (9), we get

$$r_i^*(\tau) = c_i \exp\left(-\sum_{j=1,j\neq i}^{N} \int_0^\tau \beta_{ij}(s)\mathrm{d}s\right) = c_i(1-Q_i(\tau)). \tag{12}$$

Using Eqs. (11) and (9), we obtain the $c_i=\Sigma_{j=1,j\neq i}^{N} c_j p_{ji}$, which implies that $c_i=\pi_i/C$. Here π_i is the stationary distribution of the eMC satisfying $\pi_i=\Sigma_{j=1,j\neq i}^{N}\pi_j p_{ji}$ and $\Sigma_{i=1}^{N}\pi_i=1$. Since the eMC is irreducible, the stationary distribution $\{\pi_i\}$ exists and is unique. Because the total probability must be one, the equation for C is $1=\Sigma_{i=1}^{N}\int_0^\infty r_i^*(\tau)\mathrm{d}\tau=\Sigma_{i=1}^{N}\pi_i\theta_i/C$, where θ_i is the average waiting time in state i: $\theta_i=\int_0^\infty(1-Q_i(\tau))\mathrm{d}\tau=\int_0^\infty\tau\Sigma_{j=1,j\neq i}^{N} q_{ij}(\tau)\mathrm{d}\tau$. Therefore, we obtain the stationary age structure and distribution for the sMP,

$$r_i^*(\tau) = \frac{\pi_i}{\sum_{j=1}^{N}\pi_j\theta_j}(1-Q_i(\tau)), \quad R_i^* = \int_0^\infty r_i^*(\tau)\mathrm{d}\tau = \frac{\pi_i\theta_i}{\sum_{j=1}^{N}\pi_j\theta_j}. \tag{13}$$

This result can be intuitively understood as follows: the π_i is the relative frequency for the state i being visited and θ_i is the mean dwell time of each visit. Thus, their product is proportional to the total time that the sMP is in state i, which is proportional to the probability of state i.

D. Detailed balance and time reversibility

In a stationary process, the probability of each state is time invariant and the total probability influx into the state exactly balances the total eflux out of the state. However, the stationarity does not require the influx and eflux to be balanced for each pair of states. Detailed balance, hence, is an additional constraint on a stationary stochastic process.

Definition 4: (Detailed balance) *A stationary sMP is said to have detailed balance if the net probability flux from i to j is zero for all pairs of two different states i and j*,

$$\text{net}(i \to j) \equiv \int_0^\infty \beta_{ij}(s) r_i^*(s) \mathrm{d}s - \int_0^\infty \beta_{ji}(s) r_j^*(s) \mathrm{d}s = 0, \tag{14}$$

where $\{r_1^*(\tau), r_2^*(\tau), \ldots, r_N^*(\tau)\}$ *is the stationary age distribution.*

Definition 5: (Time reversibility) *A stationary sMP is said to be time reversible if*

$$\Pr\{X_{m+k} = i_k, \Delta T_{m+k} \leqslant t_k, k = 0, 1, \ldots, n\} = \Pr\{X_{m+n-k} = i_k, \Delta T_{m+n-k} \leqslant t_k, k = 0, 1, \ldots, n\},$$

where $\Delta T_n = T_{n+1} - T_n$. *Basically, the time reversibility means that we shall see the same statistical behavior whether we read the stochastic process* $\{Y_t = x_n, t \in [T_n, T_{n+1})\}$ *forward or backward.*

Definition 6: (Direction-time independence) *A sMP is said to satisfy the direction-time independence if*

$$Q(i,j,t) = p_{ij} Q_i(t). \tag{15}$$

We have two theorems below relating detailed balance, time reversibility, and direction-time independence.

Theorem 2: *A sMP satisfies detailed balance if and only if its eMC satisfies detailed balance.*

Proof: A sMP is characterized by state sequence and the sojourn time sequence,

$$X_0, X_1, \ldots, X_n, \ldots,$$

$$\Delta T_0, \Delta T_1, \ldots, \Delta T_n, \ldots,$$

where $\Delta T_k = T_{k+1} - T_k$ is the sojourn time at state X_k. The associated eMC is characterized by only the state sequence. Consider the stationary state of the sMP. Let $\text{flux}(i \to j)$ denotes the probability flux from state i to state j. The net probability flux from state i to state j is $\text{net}(i \to j) = \text{flux}(i \to j) - \text{flux}(j \to i)$. In terms of time average, $\text{flux}(i \to j)$ is the number of $(i \to j)$ jumps of an individual system in time interval $[0, T]$ divided by T as T goes to infinity. Mathematically, it is

$$\text{flux}(i \to j) = \lim_{T \to \infty} \frac{\text{number of}(i \to j)\text{jumps in}[0,T]}{T}. \tag{16}$$

The mean waiting time and the stationary probability at state i of the eMC also can be expressed in terms of time average as

$$\pi_i = \lim_{M \to \infty} \frac{\sum_{0 \leqslant k < M, X_k = i} 1}{M}, \quad \theta_i = \lim_{M \to \infty} \frac{\sum_{0 \leqslant k < M, X_k = i} \Delta T_k}{\sum_{0 \leqslant k < M, X_k = i} 1}. \tag{17}$$

Let us consider two quantities, θ and $\Pr(i \to j)$

$$\theta \equiv \lim_{M \to \infty} \frac{1}{M} \sum_{k=0}^{M-1} \Delta T_k = \lim_{M \to \infty} \frac{1}{M} \sum_{i=1}^{N} \left(\sum_{0 \leqslant k < M, X_k = i} \Delta T_k \right),$$

$$= \sum_{j=1}^{N} \lim_{M \to \infty} \left(\frac{\sum_{0 \leqslant k < M, X_k = i} \Delta T_k}{\sum_{0 \leqslant k < M, X_k = i} 1} \frac{\sum_{0 \leqslant k < M, X_k = i} 1}{M} \right) = \sum_{j=1}^{N} \theta_i \pi_i. \tag{18}$$

θ is the overall average waiting time for a jump in the sMP.

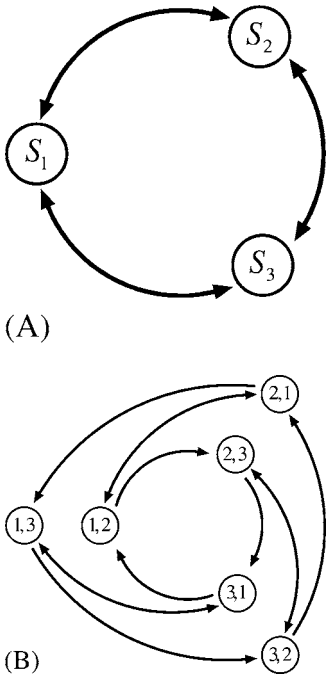

FIG. 1. (A) a sMP. (B) The new sMP obtained by transforming the sMP in (A) to satisfy the direction-time independence (Ref. 19). Note that detailed balance is destroyed in the new sMP. See the text for discussion.

$$\Pr(i \to j) \equiv \lim_{M\to\infty} \frac{\text{number of}(i \to j)\text{jumps in}\{X_0, X_1, \ldots, X_M\}}{M}, \tag{19}$$

$\Pr(i \to j)$ is the relative frequency of $(i \to j)$ jumps in a long sequence of jumps in the eMC. Using Eqs. (18) and (19), we write $\text{flux}(i \to j)$ as

$$\text{flux}(i \to j) = \lim_{M\to\infty} \frac{\text{number of}(i \to j)\text{jumps in}\{X_0, X_1, \ldots, X_M\}}{\sum_{k=0}^{M-1} \Delta T_k}$$

$$= \lim_{M\to\infty} \frac{M}{\sum_{k=0}^{M-1} \Delta T_k} \frac{\text{number of}(i \to j)\text{jumps in}\{X_0, X_1, \ldots, X_M\}}{M} = \frac{1}{\theta}\Pr(i \to j). \tag{20}$$

In terms of ensemble average, $\Pr(i \to j)$ is the relative frequency of $(i \to j)$ jumps in the ensemble of jumps occurring in a discrete time step of the eMC: $\Pr(i \to j) = \pi_i p_{ij}$. Combining this result with Eqs. (16) and (20), we arrive at

$$\text{net}(i \to j) = \frac{1}{\theta}(\pi_i p_{ij} - \pi_j p_{ji}). \tag{21}$$

Therefore, $\text{net}(i \to j) = 0$ is equivalent to $\pi_i p_{ij} - \pi_j p_{ji} = 0$.

Theorem 3: *(Chari, 1994) A sufficient and necessary condition for a stationary sMP to be time reversible is that it satisfies the direction-time independence and detailed balance.*

This theorem was proved in. Ref. 3.

Note that detailed balance alone is not enough to guarantee the time reversibility. This is illustrated in the example below. Consider a sMP in which three states form a loop, as shown in Fig. 1(a). We set

$$Q(1,2,t) = Q(2,3,t) = Q(3,1,t) = \frac{1}{2}F_+(t), \quad Q(2,1,t) = Q(3,2,t) = Q(1,3,t) = \frac{1}{2}F_-(t),$$

where $F_+(t)=\min(t,1)$ and $F_-(t)=F_+(t-1)$. The eMC has transition probabilities

$$p_{12} = p_{23} = p_{31} = \frac{1}{2}, \quad p_{21} = p_{32} = p_{13} = \frac{1}{2}.$$

The loop law is satisfied. So detailed balance is satisfied. But it is not time reversible because the waiting time distribution at state 2 in the path $1 \to 2 \to 3$ is different from that in the reversed path $3 \to 2 \to 1$.

A procedure was described in Ref. 19 to transform a sMP into a new sMP with direction-time independence. Basically, each state in the new sMP is specified by a pair of indices (i,j): i is the current state and j is the destiny state in the original sMP. In other words, each state in the new sMP specifies both the current state and the destiny state in the original sMP. If in the original sMP state j has destiny states $l_1,\ldots,l_n$, then in the new sMP, state (i,j) has destiny states $(j,l_1),\ldots,(j,l_n)$. In the new sMP, the semi-Markov kernel is $Q^{(\text{new})}((i,j),(j,l),t)=p_{jl}p_{ij}^{-1}Q(i,j,t)$, where $Q(i,j,t)$ is the semi-Markov kernel for the original sMP. As a result, the new sMP always satisfies the direction-time independence. Notice that in the new sMP (j,l) is always a destiny of (i,j). But (i,j) is a destiny of (j,l) if and only if $l=i$. As shown in Fig. 1(b), the new sMP can jump from (1,2) to (2,3) but cannot jump from (2,3) back to (1,2). Thus, the new sMP always violates detailed balance regardless of whether or not the original sMP obeys detailed balance. The transformation, while guarantees direction-time independence, simultaneously breaks down detailed balance. Therefore, the transformation cannot be used to determine the time reversibility of a sMP; it does imply that the two conditions for sMP reversibility are related.

E. Energy function and Boltzmann's law

One of the important connections between Markov processes with detailed balance and statistical mechanics of closed systems is Boltzmann's law, which states that the free energy difference between two Markov states (G_i-G_j) is related to the equilibrium probabilities as

$$\exp[-(G_i - G_j)] = \frac{R_i^*}{R_j^*} = \frac{\pi_i\theta_i}{\pi_j\theta_j} = \frac{p_{ji}}{p_{ij}}\frac{\theta_i}{\theta_j}. \tag{22}$$

where θ_i is the average waiting time in state i. If the free energies are not given, then Eq. (22) can be used to define the free energy difference between a pair of states i and j for a Markov process,

$$\Delta G_{ij} \equiv \ln\left(\frac{p_{ij}}{p_{ji}}\right) - \ln\left(\frac{\theta_i}{\theta_j}\right). \tag{23}$$

Let θ_{ij} be the conditional average waiting time in state i given that the destiny is state j. For Markov processes, θ_{ij} is independent of j and is the same as θ_i. However, for sMP, in general, we have $\theta_i \neq \theta_{ij}$. Motivated by this observation for Markov processes, we can introduce, for sMP, the free energy difference between a pair of states i and j by replacing θ_i in Eq. (23) with θ_{ij}. Note that, for a sMP, within a state there is a continuous distribution of the age. We can extend the definition of free energy to state i with age τ. Let

$$\Theta_{ij}(\tau) \equiv \Pr\{\text{staying in state } i \text{ beyond age } \tau \text{ given that the destiny is state } j\}.$$

Mathematically, $\Theta_{ij}(\tau)$ is given by $\Theta_{ij}(\tau)=1-p_{ij}^{-1}Q(i,j,\tau)$ and is related to θ_{ij} by $\theta_{ij}=\int_0^\infty \Theta_{ij}(\tau)\mathrm{d}\tau$. For sMP, it is natural to introduce the free energy difference by replacing θ_i in Eq. (23) with $\Theta_{ij}(\tau)$.

Definition 7: (Free energy) *For a sMP, the free energy difference between a pair of connected states, i with age τ and j with age τ' is*

$$\Delta G_{ij}(\tau,\tau') \equiv \ln\left(\frac{p_{ij}}{p_{ji}}\right) - \ln\left(\frac{1-p_{ij}^{-1}Q(i,j,\tau)}{1-p_{ji}^{-1}Q(j,i,\tau')}\right), \tag{24}$$

where $p_{ij}^{-1}Q(i,j,\tau)$ *is the conditional cumulative distribution of waiting time in state* i *given that the destiny is state* j. *The free energy difference between* i *and* j, *including all ages, is*

$$\Delta G_{ij} \equiv \ln\left(\frac{p_{ij}}{p_{ji}}\right) - \ln\left(\frac{\theta_{ij}}{\theta_{ji}}\right), \tag{25}$$

where $\theta_{ij}=\int_0^\infty(1-p_{ij}^{-1}Q(i,j,\tau))\mathrm{d}\tau$ *is the conditional average waiting time in state* i *given that the destiny is state* j. *In Eqs. (24) and (25), the first term is called internal energy and the second term entropy*.

Theorem 4: *A sufficient and necessary condition for the time-reversibility of a sMP is that the stationary distribution satisfies the Boltzmann's relation*

$$\frac{r_i^*(\tau)}{r_j^*(\tau')} = \exp[-\Delta G_{ij}(\tau,\tau')]. \tag{26}$$

Proof of Theorem 4 is presented in Appendix A.

F. Symmetric operator and time reversibility

The differential equation in Eq. (10) can be rewritten in an integral form,

$$\int_0^\infty \mathrm{d}s\left\{\frac{\partial}{\partial t}r_i(s,t) + \sum_{j=1,j\neq i}^{N}[\beta_{ij}(s)r_i(s,t) - \beta_{ji}(s)r_j(s,t)]\right\} = 0, \tag{27}$$

which very much resembles the forward equation for a Markov process. If the sMP is a Markov process, then $\beta_{ij}(s)=k_{ij}$ independent of s, and $R_i(t)\equiv\int_0^\infty r_i(s,t)\mathrm{d}s$ satisfies a master equation

$$\frac{\mathrm{d}}{\mathrm{d}t}R_i(t) = \sum_{j=1,j\neq i}^{N}[k_{ji}R_j(t) - k_{ij}R_i(t)]. \tag{28}$$

We consider a linear operator $\mathbf{L}$ on $\mathbf{V}=(V_1,V_2,\ldots,V_N)$ defined as $(\mathbf{LV})_i\equiv\Sigma_{j=1,j\neq i}^N[k_{ji}V_j-k_{ij}V_i]$. It can be shown that the Markov process is time reversible if and only if the linear operator $\mathbf{L}$ is symmetric with respect to the inner product defined as $\langle\mathbf{U},\mathbf{V}\rangle\equiv\Sigma_{i=1}^N U_iV_i/R_i^*$, where R_i^* is the stationary distribution of the Markov process. Motivated by this observation for Markov processes, we consider its counterpart for sMP. For a Markov process, we notice that $k_{ij}=Q(i,j,\tau)/(\theta_iQ_i(\tau))$, where θ_i is the mean waiting time at state i, and that R_i^* is proportional to $\pi_i\theta_i$. Thus, for a sMP, we have the definition and the theorem below relating the symmetry of an operator and the time reversibility.

Definition 8: *Consider asMP with kernel functions* $Q(i,j,t)$. *Let us define a linear operator* $\mathcal{L}$ *on* $\mathbf{v}(\tau)=(v_1,v_2,\cdots,v_N)(\tau)$ *defined as*

$$(\mathcal{L}\mathbf{v})_i(\tau) \equiv \sum_{j=1,j\neq i}^{N}\left[\frac{Q(j,i,\tau)}{\theta_jQ_j(\tau)}v_j(\tau) - \frac{Q(i,j,\tau)}{\theta_iQ_i(\tau)}v_i(\tau)\right]. \tag{29}$$

We call it the linear operator associated with the sMP.

Theorem 5: *A sufficient and necessary condition for the time reversibility of a sMP is that its associated linear operator* $\mathcal{L}$ *is symmetric with respect to the inner product*

$$\langle\mathbf{u},\mathbf{v}\rangle \equiv \int_0^\infty\int_0^\infty\sum_{i=1}^{N}u_i(\tau)v_i(\tau')\frac{1}{\pi_i\theta_i}\mathrm{d}\pi\mathrm{d}\tau'. \tag{30}$$

That is, $\langle\mathbf{u},\mathcal{L}\mathbf{v}\rangle=\langle\mathcal{L}\mathbf{u},\mathbf{v}\rangle$ *for all* $\mathbf{u}$ *and* $\mathbf{v}$.

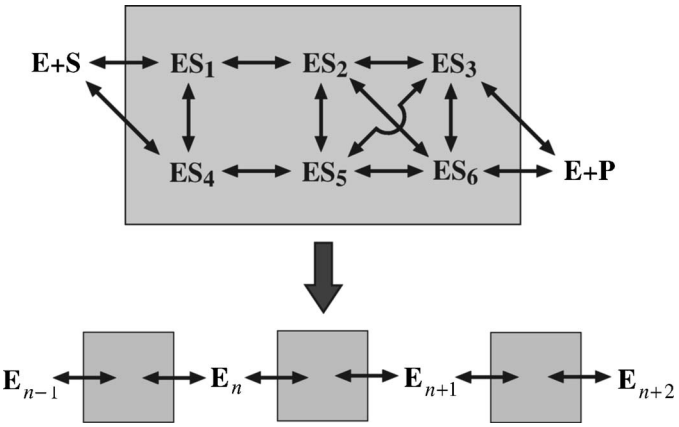

FIG. 2. Modeling enzyme reactions with Markov processes and semi-Markov processes. The top panel shows one cycle of an enzyme reaction modeled as a Markov process. The bottom panel shows a sequence of cycles modeled as a semi-Markov process in which each cycle is one semi-Markov step. In the bottom panel, the subscript denotes the number of cycles completed. Only three cycles are shown.

Proof of Theorem 5 is presented in Appendix B.

III. SINGLE-MOLECULE ENZYME KINETICS AND THE RISE OF DIRECTION-TIME INDEPENDENCE

In this section we examine an enzyme kinetic problem. The goal is to illustrate how the direction-time independence arises in connection to detailed balance in a Markov model. More specifically, we are going to show that if the Markov model is detailed balanced, then the conditional cycle time distribution for a forward enzymatic cycle is the same as that for a backward enzymatic cycle. In other words, if one reaction cycle is modeled as one semi-Markov step, then the sMP satisfies the direction-time independence. We also present an example to show that when the Markov model is not detailed balanced, then the waiting time distributions can be different for the forward and backward cycles.

A. Enzyme reactions with well defined starting/ending state for each cycle

Let us consider a general reversible enzymatic reaction with a single enzyme molecule. The enzyme molecule is a catalyst which aids in converting reactants to products, one at a time, and vice versa. In biochemical contexts, it is assumed that the change in the amount of reactants and products in the surrounding media is negligible. Hence, the system is renewal after either a forward cycle or a backward cycle. The top panel of Fig. 2 shows an enzyme reaction modeled as a Markov process. The enzyme reaction shown has a well defined starting/ending state while the enzyme molecule returns to exactly its starting state. If we focus on the starting and ending states of each cycle, then the sequence of cycles becomes a semi-Markov process in which each cycle is one semi-Markov step, as shown in the bottom panel of Fig. 2. We now show that if the Markov process describing one cycle satisfies detailed balance, then the conditional waiting time distribution for forward cycles is the same as that for backward cycles. In other words, the semi-Markov process shown in the bottom panel of Fig. 2 satisfies the direction-time independence.

To prove this result, it is sufficient to consider two consecutive cycles described by the Markov process. For mathematical convenience, we number the states in the two cycles from 0 to $2N$, where N is the number of states in each cycle. Figure 3 illustrates two cycles of the reaction in Fig. 2, with states appropriately numbered. S_N is the starting/ending state that separates two cycles. We study the waiting time distributions for forward and backward cycles. Suppose the system arrives at $t=0$ at state S_N, which separates two cycles. We make states S_0 and S_{2N} absorb-

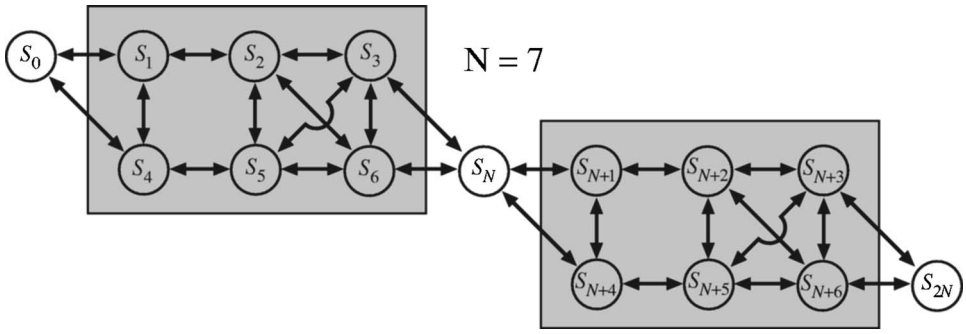

FIG. 3. Two cycles of the reaction in Fig. 2, with states numbered. In the reaction shown, the number of states in one cycle is $N=7$. State S_N is the ending state of the cycle on the left and is the starting state of the cycle on the right.

ing. We assume the Markov process satisfies detailed balance. More specifically, we assume each state has a free energy G_i. The transition rates in the Markov process and the free energies satisfy detailed balance

$$\frac{k_{i\to j}}{k_{j\to i}} = \exp\left(\frac{G_i - G_j}{k_B T}\right), \tag{31}$$

where $G_{N+j}-G_j=\Delta G$ and ΔG is the free energy charge of one reaction cycle.

Theorem 6: *Suppose the system starts at $t=0$ at S_N. Suppose the Markov process satisfies detailed balance [Eq. (31)]. Let $\rho_+(t)$ be the probability density of the waiting time of escaping to S_{2N} (forward cycle) and $\rho_-(t)$ be the probability density of the waiting time of escaping to S_0 (backward cycle). We have*

$$\rho_-(t) = \rho_+(t)\exp\left(\frac{\Delta G}{k_B T}\right). \tag{32}$$

Proof: Let k_i be the total rate of jumping out of state i: $k_i \equiv \Sigma_{j\neq i} k_{i\to j}$, where $k_{i\to j}$ is the transition rate from state i to state j in the Markov process. So $k_{i\to j}/k_i$ is the probability of jumping from state i to state j. Let $\rho(t;r_1,r_2,\ldots,r_m)$ denotes the probability density of the sum of m independent exponential random variables with parameters r_1, $r_2,\ldots,r_m$. From the definition, we see that $\rho(t;r_1,r_2,\ldots,r_m)$ satisfies

$$\rho(t;r_1,r_2,\ldots,r_m) = \rho(t;r_{i_1},r_{i_2},\ldots,r_{i_m}), \tag{33}$$

where $(i_1,i_2,\ldots,i_m)$ is a permutation of $(1,2,\ldots,m)$. This property is very important in our analysis below.

Expressing $\rho_+(t)$ and $\rho_-(t)$ in terms of an infinite sum of conditional waiting time distributions along all possible escape paths weighted by the probability of taking each escape path, we have

$$\rho_+(t) = \sum_{\{F_1F_2\cdots F_{m-1}F_m\}} \left[\rho(t;k_{F_1},k_{F_2},\ldots,k_{F_{m-1}}) \prod_{j=1}^{m-1}\left(\frac{k_{F_j\to F_{j+1}}}{k_{F_j}}\right)\right], \tag{34}$$

$$\rho_-(t) = \sum_{\{B_1B_2\cdots B_{m-1}B_m\}} \left[\rho(t;k_{B_1},k_{B_2},\ldots,k_{B_{m-1}}) \prod_{j=1}^{m-1}\left(\frac{k_{B_j\to B_{j+1}}}{k_{B_j}}\right)\right]. \tag{35}$$

In Eq. (34), the factor $\Pi_{j=1}^{m-1}(k_{F_j\to F_{j+1}}/k_{F_j})$ is the probability of taking the escape path $\{F_1F_2\cdots F_{m-1}F_m\}$ and the factor $\rho(t;k_{F_1},k_{F_2},\ldots,k_{F_{m-1}})$ is the conditional probability density of waiting time along the escape path $\{F_1F_2\cdots F_{m-1}F_m\}$. The two factors in Eq. (35) have the similar meanings. In Eq. (34) the summation is taken over all possible forward escape paths and in Eq.

(35) the summation is taken over all backward escape paths. More precisely, $\{F_1F_2\cdots F_{m-1}F_m\}$ is called a forward escape path if it satisfies

$$F_1 = N, \quad \text{none of } F_2, \ldots, F_{m-1} \text{ is } 0 \text{ or } 2N, \quad F_m = 2N, \tag{36}$$

$\{B_1B_2\cdots B_{m-1}B_m\}$ is called a backward escape path if it satisfies

$$B_1 = N, \quad \text{none of } B_2, \ldots, B_{m-1} \text{ is } 0 \text{ or } 2N, \quad B_m = 0. \tag{37}$$

There is a one-to-one correspondence between forward escape paths and backward escape paths. Suppose $\{F_1F_2\cdots F_{m-1}F_m\}$ is a forward escape path. Starting at S_N at $t=0$, the path may come back to S_N a number of times before finally escaping to S_{2N} without touching S_N again. Let l be the index along the path where the path goes through S_N for the last time before escaping to S_{2N}. We construct backward escape path $\{B_1B_2\cdots B_{m-1}B_m\}$ corresponding to forward escape path $\{F_1F_2\cdots F_{m-1}F_m\}$ as follows:

$$B_j = F_j, \quad 1 \leqslant j \leqslant l$$

$$B_j = F_{m+l-j} - N, \quad l+1 \leqslant j \leqslant m. \tag{38}$$

It is straightforward to verify that $\{B_1B_2\cdots B_{m-1}B_m\}$ defined in Eq. (38) satisfies condition (37), and thus, is a backward escape path. Basically, we keep the segment of the path from index 1 to index l. For the segment of the path from index $(l+1)$ to m, we reverse it and shift it back by one cycle. For example, in the diagram shown in Fig. 3, $\{S_N\ S_{N+1}\ S_{N+4}\ S_N\ S_{N+1}\ S_{N+2}\ S_{N+6}\ S_{N+3}\ S_{2N}\}$ is a forward escape path. The corresponding backward escape path is $\{S_N\ S_{N+1}\ S_{N+4}\ S_N\ S_3\ S_6\ S_2\ S_1\ S_0\}$. For the pair of corresponding forward and backward escape paths $\{F_1F_2\cdots F_{m-1}F_m\}$ and $\{B_1B_2\cdots B_{m-1}B_m\}$, let us examine the terms in Eqs. (34) and (35). Since all cycles are described by the same Markov process, we have $k_{(N+i)\to(N+j)}=k_{i\to j}$ and $k_{N+i}=k_i$. Combining this with property (33), we obtain

$$\frac{\rho(t;k_{B_1},k_{B_2},\ldots,k_{B_{m-1}})}{\prod_{j=1}^{m-1} k_{B_j}} = \frac{\rho(t;k_{F_1},k_{F_2},\ldots,k_{F_{m-1}})}{\prod_{j=1}^{m-1} k_{F_j}}. \tag{39}$$

Using the correspondence (38) between the forward and backward escape paths, we have

$$\prod_{j=1}^{m-1} k_{B_j\to B_{j+1}} = \left(\prod_{j=1}^{l-1} k_{B_j\to B_{j+1}}\right)\left(\prod_{j=l}^{m-1} k_{B_j\to B_{j+1}}\right),$$

$$= \left(\prod_{j=1}^{l-1} k_{F_j\to F_{j+1}}\right)\left(\prod_{j=l}^{m-1} k_{(F_{m+l-j})\to(F_{m+l-1-j})}\right).$$

Using index substitution $i=m+l-1-j$ and using detailed balance (31), we obtain

$$\prod_{j=1}^{m-1} k_{B_j\to B_{j+1}} = \left(\prod_{j=1}^{l-1} k_{F_j\to F_{j+1}}\right)\left(\prod_{i=l}^{m-1} k_{F_{i+1}\to F_i}\right)$$

$$= \left(\prod_{j=1}^{l-1} k_{F_j\to F_{j+1}}\right)\left(\prod_{j=l}^{m-1} k_{F_j\to F_{j+1}}\right)\exp\left(\frac{\sum_{j=l}^{m-1}(G_{F_{j+1}} - G_{F_j})}{k_BT}\right)$$

$$= \left(\prod_{j=1}^{m-1} k_{F_j\to F_{j+1}}\right)\exp\left(\frac{\Delta G}{k_BT}\right). \tag{40}$$

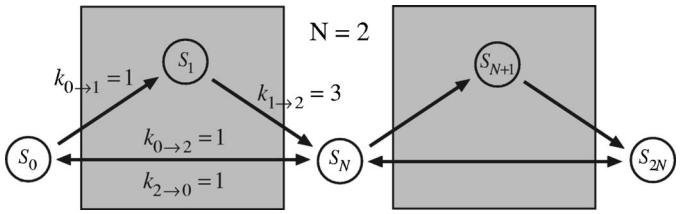

FIG. 4. Two cycles of an enzymatic reaction. Transition directions and rates are marked in the figure. The reaction shown does not obey detailed balance.

Since Eqs. (39) and (40) are true for every pair of forward and backward escape paths, we conclude

$$\rho_-(t) = \rho_+(t)\exp\left(\frac{\Delta G}{k_B T}\right).$$

For escape problems in reliability studies, there is a general closed form solution.[19] However, for the escape problem shown in Fig. 3, the closed form solution does not readily lead to conclusion (32). Our approach of proving the conclusion shows how detailed balance is used and why it is important.

B. Breakdown of direction-time independence when detailed balance is violated

To demonstrate that detailed balance in the Markov process is necessary for the direction-time independence in the sMP derived from the Markov process, we consider an enzymatic reaction described by the Markov process shown in Fig. 4. Only two cycles are shown. Each cycle has $N=2$ states. We set

$$k_{0\to2} = k_{2\to0} = 1, \quad k_{0\to1} = 1, \quad k_{1\to2} = 3, \quad k_{1\to0} = k_{2\to1} = 0$$

It follows that $k_0=3$ and $k_1=1$. All cycles are described by the same Markov process: $k_{N+i\to N+j} = k_{i\to j}$ and $k_{N+i}=k_i$. It is straightforward to verify that the Markov process violates detailed balance,

$$\frac{k_{0\to2}k_{2\to1}k_{1\to0}}{k_{0\to1}k_{1\to2}k_{2\to0}} = 0 \neq 1. \tag{41}$$

Starting at state N, there are two forward escape paths: $\{N,2N\}$ and $\{N,N+1,2N\}$ and one backward escape path $\{N,0\}$. Notice that the lack of one-to-one correspondence between the forward and backward escape paths is caused by the violation of detailed balance. Using Eqs. (34) and (35), we obtain

$$\rho_+(t) = \exp(-3t) + 3t\exp(-3t), \quad \rho_-(t) = \exp(-3t).$$

It is clear that $\rho_+(t)$ is not proportional to $\rho_-(t)$. As a result, in the sMP describing one cycle as one semi-Markov step, the direction-time independence is not satisfied. This is caused by the fact that in the Markov process governing the enzymatic transition steps within a cycle in Fig. 4, detailed balance is violated.

C. Aggregates of Markov states as a semi-Markov state and its direction-time independence

An aggregate is a group of states. For more background of the subject see.[5,8] Let us consider a group of states in a Markov process. Let us denote the group by G. We assume group G is not empty and does not contain all states of the Markov process. Below "j is directly connected from i" means $k_{i\to j}>0$. "j is directly connected to i" means $k_{j\to i}>0$. We introduce three categories of states relative to group G.

- The outer boundary of G denoted by $\partial_{\text{out}}G$: we say that state j is in the outer boundary of G if j is not in G and j is directly connected from a state in G.
- The inner boundary of G denoted by $\partial_{\text{in}}G$: we say that state j is in the inner boundary of G if j is in G and j is directly connected to a state in $\partial_{\text{out}}G$.
- Internal part of G denoted by $G\backslash\partial_{\text{in}}G$: we say that j is in the internal part of G if j is in G and j is not in the inner boundary of G

We consider the escape problem of starting at state $i \in \partial_{\text{in}}G$ at time 0 with all states in $\partial_{\text{out}}G$ are absorbing. Let $q_j(t;i)$ denotes the probability density of starting at state $i \in \partial_{\text{in}}G$ at time 0 and escaping to state $j \in \partial_{\text{out}}G$ at time t. Group G can be considered a semi-Markov state if $q_j(t;i)$ is independent of i. That is, the probability density of the escape time is independent of where the system enters group G. By definition, $\partial_{\text{in}}G$ is not empty. If $\partial_{\text{in}}G$ has only one state, then $q_j(t;i)$ is independent of i. In this case the direction-time independence is also satisfied. If $\partial_{\text{in}}G$ has two or more states, then we have several results.

1. If $q_j(t;i)$ is independent of i, then every state $i \in \partial_{\text{in}}G$ must be directly connected to every state $j \in \partial_{\text{out}}G$.

   Proof: We prove it by contradiction. Suppose state $i_1 \in \partial_{\text{in}}G$ is not directly connected to state $j \in \partial_{\text{out}}G$. By the definitions of $\partial_{\text{in}}G$ and $\partial_{\text{out}}G$, we see that there exists state $i_2 \in \partial_{\text{in}}G$ that is directly connected to state j. We have

   $$q_j(0;i_1) = 0, \quad q_j(0;i_2) = k_{i_2 \to j} > 0,$$

   which contradicts with that $q_j(t;i)$ is independent of $i \in \partial_{\text{in}}G$.
2. If $q_j(t;i)$ is independent of i, then we have

   $$k_{i_1 \to j} = k_{i_2 \to j}, \quad \text{for all } i_1 \in \partial_{\text{in}}G, i_2 \in \partial_{\text{in}}G, \quad j \in \partial_{\text{out}}G.$$

   Proof: This follows immediately from $q_j(0;i) = k_{i \to j}$.
3. If $q_j(t;i)$ is independent of i, then the direction-time independence is satisfied.

   Proof: By the definitions of $\partial_{\text{in}}G$ and $\partial_{\text{out}}G$, the system can escape to $\partial_{\text{out}}G$ only from a state in $\partial_{\text{in}}G$. Suppose the system starts in state $i_1 \in \partial_{\text{in}}G$. Since group G is a semi-Markov state it does not matter in which state of $\partial_{\text{in}}G$ the system starts. Suppose $(i_1, i_2, \ldots, i_m, j_1)$ is an escape path to $j_1 \in \partial_{\text{out}}G$, where $i_m \in \partial_{\text{in}}G$. Then $(i_1, i_2, \ldots, i_m, j_2)$ is an escape path to $j_2 \in \partial_{\text{out}}G$. This one-to-one correspondence tells us that the conditional distribution of escape time to state $j_1 \in \partial_{\text{out}}G$ is the same as the conditional distribution of escape time to state $j_2 \in \partial_{\text{out}}G$. The probability of escaping to state $j \in \partial_{\text{out}}G$ is given by $p_j = k_{i_m \to j} / (\Sigma_{l \in \partial_{\text{out}}G} k_{i_m \to l})$ which, as we showed in item 2 above, is independent of i_m.

Results listed above indicate that, in general, an aggregation of Markov states is not a semi-Markov state unless every states in the inner boundary of the aggregation is directly connected to every state in the outer boundary of the aggregation (this condition is automatically satisfied if the aggregation has one and only one state in the inner boundary). This conclusion seems to contradict with the example shown in Fig. 3. It is important to point out that the semi-Markov process shown in the bottom panel of Fig. 3 is *not* obtained by aggregating Markov states. In the bottom panel of Fig. 3, the subscript n means the number of cycles completed. So n is well defined for states S_0, S_N, and S_{2N}. For states between S_0 and S_N, n is not completely determined by the state. Rather n depends on whether or not the state is reached from S_0 or from S_N. In Fig. 4, if state S_4 is reached from S_N, then S_4 has the same value of n as S_N. If S_4 is reached from S_0, then S_4 has the same value of n as S_0. This may sound very complicated, confusing or even contradicting. A mathematically simple and biochemically practical way of tracking the value of n is the follows. Start the system at S_0 and start with $n=0$. When a reactant binding is followed by a product release, the value of n is increased by 1. A reactant binding followed by a reactant release does not change the value of n. Similarly when a product binding is followed by a reactant release, the value of n is decreased

by 1. A product binding followed by a product release does not change the value of n. This is how the semi-Markov process shown in the bottom panel of Fig. 3 is constructed.

ACKNOWLEDGMENTS

H. Wang was partially supported by NSF.

APPENDIX A: PROOF OF THEOREM 4

Part 1: We show that the time reversibility implies Boltzmann's relation. The time reversibility implies detailed balance $\pi_i p_{ij}=\pi_j p_{ji}$ and the direction-time independence $Q(i,j,t)=p_{ij}Q_i(t)$. The free energy between a pair of connected states, i with age τ and j with age τ' is

$$G_{ij}(\tau,\tau')=\ln\left(\frac{p_{ij}}{p_{ji}}\right)-\ln\left(\frac{1-Q_i(\tau)}{1-Q_j(\tau')}\right). \tag{A1}$$

Substituting Eqs. (A1) and (13) into Eq. (26), and using detailed balance $\pi_i p_{ij}=\pi_j p_{ji}$, it is straightforward to verify that the Boltzmann relation is satisfied.

Part 2: We show that Boltzmann's relation (26) implies the time reversibility. We only need to show that Boltzmann's relation implies both detailed balance and the direction-time independence.

We first derive detailed balance. Setting $\tau=0$ and $\tau'=0$ in Boltzmann's relation (26), using Eq. (13) and the fact that $Q(i,j,0)=0$, we get $\pi_i/\pi_j=p_{ji}/p_{ij}$ which is equivalent to detailed balance for the eMC.

To derive the direction-time independence, we let τ go free and keep $\tau'=0$ in Boltzmann's relation (26). Using detailed balance we just derived, we obtain $Q_i(\tau)=p_{ij}^{-1}Q(i,j,\tau)$ for all τ and j, which leads to the direction-time independence.

APPENDIX B: PROOF OF THEOREM 5

We start by rewriting $\langle \mathbf{u},\mathcal{L}\mathbf{v}\rangle$,

$$\begin{aligned}\langle \mathbf{u},\mathcal{L}\mathbf{v}\rangle &= \int_0^\infty\int_0^\infty \sum_{i=1}^N u_i(\tau)(\mathcal{L}\mathbf{v})_i(\tau')\frac{1}{\pi_i\theta_i}\mathrm{d}\tau\mathrm{d}\tau' = \int_0^\infty\int_0^\infty \sum_{i=1}^N\sum_{j=1,j\neq i}^N u_i(\tau)\left[\frac{Q(j,i,\tau')}{\theta_j Q_j(\tau')}v_j(\tau')\right.\\ &\left.-\frac{Q(i,j,\tau')}{\theta_i Q_i(\tau')}v_i(\tau')\right]\frac{1}{\pi_i\theta_i}\mathrm{d}\tau\mathrm{d}\tau' = \int_0^\infty\int_0^\infty \sum_{j=1}^N\sum_{i=1,i\neq j}^N u_i(\tau)\frac{1}{\pi_i\theta_i\theta_j}\frac{Q(j,i,\tau')}{Q_j(\tau')}v_j(\tau')\mathrm{d}\tau\mathrm{d}\tau'\\ &-\int_0^\infty\int_0^\infty \sum_{i=1}^N u_i(\tau)\frac{1}{\pi_i\theta_i^2}v_i(\tau')\mathrm{d}\tau\mathrm{d}\tau'.\end{aligned} \tag{B1}$$

In the above we have used the fact that $\Sigma_{j=1,j\neq i}^N Q(i,j,\tau')=Q_i(\tau')$. The second part on the right hand side of Eq. (B1), corresponding to the diagonal part of the linear operator, is already in a symmetric form for all $\mathbf{u}$ and $\mathbf{v}$. Thus, $\langle \mathbf{u},\mathcal{L}\mathbf{v}\rangle=\langle \mathcal{L}\mathbf{u},\mathbf{v}\rangle$ for all $\mathbf{u}$ and $\mathbf{v}$ is equivalent to

$$\int_0^\infty\int_0^\infty u_i(\tau)\frac{1}{\pi_i\theta_i\theta_j}\frac{Q(j,i,\tau')}{Q_j(\tau')}v_j(\tau')\mathrm{d}\tau\mathrm{d}\tau' = \int_0^\infty\int_0^\infty u_i(\tau)\frac{Q(i,j,\tau)}{Q_i(\tau)}\frac{1}{\pi_j\theta_j\theta_i}v_j(\tau')\mathrm{d}\tau\mathrm{d}\tau'$$

for all functions $u_i(\tau)$ and $v_j(\tau')$ $(i\neq j)$, which is equivalent to

$$\frac{1}{\pi_i}\frac{Q(j,i,\tau')}{Q_j(\tau')}=\frac{1}{\pi_j}\frac{Q(i,j,\tau)}{Q_i(\tau)} \quad \text{for all } \tau \text{ and } \tau'. \tag{B2}$$

Part 1: We show that the time reversibility implies Eq. (B2). The time reversibility implies detailed balance $\pi_i p_{ij}=\pi_j p_{ji}$ and the direction-time independence $Q(i,j,t)=p_{ij}Q_i(t)$. Using these two, we have

$$\frac{1}{\pi_i}\frac{Q(j,i,\tau')}{Q_j(\tau')} = \frac{p_{ji}}{\pi_i} = \frac{p_{ij}}{\pi_j} = \frac{1}{\pi_j}\frac{Q(i,j,\tau)}{Q_i(\tau)} \quad \text{for all } \tau \text{ and } \tau'.$$

Part 2: We show that Eq. (B2) implies the time reversibility. We only need to show that Eq. (B2) implies both detailed balance and the direction-time independence. We first derive detailed balance. In Eq. (B2), taking the limit as $\tau \to \infty$ and $\tau' \to \infty$, and using $Q(i,j,\infty) = p_{ij}$ and $Q_i(\infty) = 1$, we obtain $p_{ji}/\pi_i = p_{ij}/\pi_j$ which leads to detailed balance for the eMC. To derive the direction-time independence, we take the limit as $\tau' \to \infty$ but leave τ free in Eq. (B2). Taking the limit of Eq. (B2), multiplying by π_j, and using detailed balance we just derived, we arrive at

$$\frac{Q(i,j,\tau)}{Q_i(\tau)} = \frac{\pi_j}{\pi_i}\lim_{\tau' \to \infty}\frac{Q(j,i,\tau')}{Q_j(\tau')} = \frac{\pi_j}{\pi_i}p_{ji} = p_{ij} \tag{B3}$$

for all τ, which is the direction-time independence.

[1] Bergmann, P. G. and Lebowitz, J. L., "New approach to nonequilibrium processes," Phys. Rev. **99**, 578–587 (1955).
[2] Callen, H. B. and Welton, T. A., "Irreversibility and generalized noise," Phys. Rev. **83**, 34–40 (1951).
[3] Chari, M. K., "On reversible semi-Markov processes," Oper. Res. Lett. **15**, 157–161 (1994).
[4] Cinlar, E., *Introduction to Stochastic Processes* (Pretice-Hall, Englewood Cliffs, NJ, 1975).
[5] Colquhoun, D. and Hawkes, A. G., "On the stochastic properties of single ion channels," Proc. R. Soc. London, Ser. B **211**, 205–235 (1981).
[6] Érdi, P. and Tóth J., *Mathematical Methods of Chemical Reactions: Theory and Applications of Deterministic and Stochastic Models* (Manchester University Press, Manchester, 1989).
[7] Feller, W., *An Introduction to Probability Theory and Its Applications*, 2nd Ed. (Wiley, New York, 1957), Vol. 1.
[8] Fredkin, D. R. and Rice, J. A., "On aggregated Markov processes," J. Appl. Probab. **23**, 208–214 (1986).
[9] Hänggi, P., Talkner, P., and Borkovec, M., "Reaction-rate theory: fifty years after Kramers," Rev. Mod. Phys. **62**, 251–341 (1990).
[10] Hill, T. L., *Free Energy Transduction in Biology* (Academic, New York, 1977).
[11] Jiang, D.-Q, Qian, M., and Qian, M.-P., *Mathematical Theory of Nonequilibrium Steady-State LNM* (Springer, New York, 2004), Vol. 1833.
[12] Kelly, F. P., *Reversibility and Stochastic Networks* (Wiley, New York, 1979).
[13] Kim, T. S. and Chae, K. C., "A note on the age of the absorbing sem-Markov process," J. Oper. Res. Soc. **48**, 843–845 (1997).
[14] Kolomeisky, A. B. and Fisher, M. E., "Extended kinetic models with waiting-time distributions: exact results," J. Chem. Phys. **113**, 10867–10877 (2000).
[15] Kolomeisky, A. B., Stukalin, E. B., and Popov, A. A., "Understanding mechanochemical coupling in kinesins using first-passage-time processes," Phys. Rev. E **71**, 031902 (2005).
[16] Kou, S. C., Cherayil, B. J., Min, W., English, B. P., and Xie, X. S., "Single-molecule Michaelis-Menten equations," J. Phys. Chem. B **109**, 19068–19081 (2005).
[17] Kramers, H. A., "Brownian motion in a field of force and the diffusion model of chemical reactions," Physica (Amsterdam) **7**, 284–304 (1940).
[18] Lewis, G. N., "A new principle of equilibrium," Proc. Natl. Acad. Sci. U.S.A. **11**, 179–83 (1925)
[19] Limnios, N. and Oprisan, G., *Semi-Markov Processes and Reliability* (Birkhäuser, Boston, 2001).
[20] Mehta, A. D., Rief, M., Spudich, J. A., Smith, D. A., and Simmons, R. M., "Single-molecule biomechanics with optical methods," Science **283**, 1689–95 (1999).
[21] McQuarrie, D. A., "Stochastic approach to chemical kinetics," J. Appl. Probab. **4**, 413–478 (1967).
[22] Nicolis, G. and Prigogine, I., *Self-Organization in Nonequilibrium Systems* (Wiley, New York, 1977).
[23] Parzen, E., *Stochastic Processes* (SIAM, Philadelphia, 1999).
[24] Qian, H., "A simple theory of motor protein kinetics and energetics," Biophys. Chem. **67**, 263–267 (1997).
[25] Qian, H., "Mathematical formalism for isothermal linear irreversibility," Proc. R. Soc. London, Ser. A **457**, 1645–1655 (2001).
[26] Qian, H., "Open-system nonequilibrium steady-state: statistical thermodynamics, fluctuations and chemical oscillations," J. Phys. Chem. B **110**, 15063–15074 (2006).
[27] Qian, H. and Elson, E. L., "Single-molecule enzymology: stochastic Michaelis-Menten kinetics," Biophys. Chem. **101**, 565–576 (2002).
[28] Qian, H. and Wang, H., "Continuous time random walks in closed and open single-molecule systems with microscopic reversibility," Europhys. Lett. **76**, 15–21 (2006).
[29] Qian, H. and Xie, X. S., "Generalized Haldane equation and fluctuation theorem in the steady state cycle kinetics of single enzymes," Phys. Rev. E **74**, 010902(R) (2006).
[30] Schnakenberg, J., "Network theory of microscopic and macroscopic behavior of master equation systems," Rev. Mod. Phys. **48**, 571–585 (1976).
[31] Tolman, R. C., "Duration of molecules in upper quantum states," Phys. Rev. **23**, 693–709 (1924).
[32] Xie, X. S. and Lu, H. P., "Single-molecule enzymology," J. Biol. Chem. **274**, 15967–15970 (1999).