Large deviations and fluctuation theorems for cycle currents defined in the loop-erased and spanning tree manners: A comparative study

Yuhao Jiang,[1,*] Bingjie Wu,[1,2,*] and Chen Jia[1,†]

[1]*Applied and Computational Mathematics Division, Beijing Computational Science Research Center, Beijing 100193, China*

[2]*LMAM, School of Mathematical Sciences, Peking University, Beijing 100871, China*

(Received 4 April 2022; accepted 1 February 2023; published 27 March 2023)

The cycle current is a crucial quantity in stochastic thermodynamics. The absolute and net cycle currents of a Markovian system can be defined in the loop-erased (LE) or spanning tree (ST) manner. Here we make a comparative study between the large deviations and fluctuation theorems for LE and ST currents, i.e., cycle currents defined in the LE and ST manners. First, we derive the exact joint distribution and large deviation rate function for the LE currents of a system with a cyclic topology and also obtain the exact rate function for the ST currents of a general system. The relationship between the rate functions for LE and ST currents is clarified and the analytical results are applied to examine the fluctuations in the product rate of a three-step reversible enzyme reaction. Furthermore, we examine various types of fluctuation theorems satisfied by LE and ST currents and clarify their ranges of applicability. We show that both the absolute and net LE currents satisfy the strong form of all types of fluctuation theorems. In contrast, the absolute ST currents do not satisfy fluctuation theorems, while the net ST currents only satisfy the weak form of fluctuation theorems under the periodic boundary condition. Finally, the fluctuation theorems for cycle currents are applied to study the fluctuations in entropy production along a single stochastic trajectory.

DOI: 10.1103/PhysRevResearch.5.013207

I. INTRODUCTION

Over the past two decades, significant progress has been made in stochastic thermodynamics [1–3], which has grown to become an influential branch of nonequilibrium statistical physics. In this field, a thermodynamic system is usually modeled by a Markov process. Markov chains, whose state spaces are discrete, are the most fundamental and important dynamic model since any Markov process can always be approximated by a Markov chain. Along this line, an equilibrium state is defined as a reversible Markov process and the deviation from equilibrium is usually quantified by the concept of entropy production, which can be represented as a bilinear function of thermodynamic fluxes and forces [4,5]. It has long been noted by Kolmogorov [6,7] that the reversibility of a Markov chain can be characterized by its cycle dynamics: the system is reversible if and only if the product of transition probabilities along each cycle and that along its reversed cycle are exactly the same, which generalizes the Wegscheider condition for detailed balanced chemical reaction networks. An incisive observation is that the entropy production can be decomposed along cycles, with the thermodynamics fluxes being the cycle currents (also called cycle fluxes or circulations) and with the thermodynamic forces being the cycle affinities [8].

The cycle representation theory of Markov chains has found wide applications in physics, chemistry, and biology [9,10]. In fact, the current of a cycle can be defined in several different ways. Two common definitions are based on the spanning tree (ST) and loop-erased (LE) methods. Hill [11–14] and Schnakenberg [8] developed a network theory and defined the currents for a family of *fundamental cycles*. In this theory, a ST is associated with the transition diagram of a Markov chain, which is a directed graph. Each edge of the graph that does not belong to the ST, which is called a chord, will generate a fundamental cycle. The current of a fundamental cycle is defined as the number of times that the associated chord is traversed per unit time. Qian and coworkers [15–17] further developed the cycle representation theory and defined the currents for all *simple cycles* of the graph, i.e., cycles with no repeated vertices except the beginning and ending vertices. In this theory, the trajectory of a Markovian system is tracked. Once a cycle is formed, it is erased from the trajectory and we further keep track of the remaining trajectory until the next cycle is formed. The current of a simple cycle then is defined as the number of times that the cycle is formed per unit time. Recently, another type of cycle current has been proposed based on the idea of sequence matching [18–20]. In this theory, the currents are defined for *all cycles* of the graph, i.e., directed paths with the first and last vertices being equal.

All types of cycle currents can also be defined along a single stochastic trajectory. One of the major advances

*These authors contributed equally to this work.

†Corresponding author: chenjia@csrc.ac.cn

in stochastic thermodynamics is the finding that a broad class of thermodynamic quantities such as entropy production and cycle currents satisfy various types of fluctuation theorems [21–32], which provide nontrivial generalizations of the second law of thermodynamics in terms of equalities rather than inequalities. For cycle currents defined in the ST manner, Andrieux and Gaspard [33] proved that the fluctuation theorem holds for *net* cycle currents in the long-time limit. Moreover, Polettini and Esposito [34] showed that the transient fluctuation theorem at any finite time holds if the definition of cycle currents is slightly modified. For the cycle currents defined in the LE manner, Andrieux and Gaspard [35] and Jia *et al.* [36] proved that all types of fluctuation theorems and symmetric relations are satisfied for both the *absolute* and *net* cycle currents. For cycle currents defined in the sequence matching manner, the corresponding fluctuation theorems and symmetric relations have also been developed recently [20]. The fluctuation theorems for cycle currents have also been developed for some stochastic processes with continuous state space, such as Langevin dynamics on circles [37].

From the mathematical perspective, another important question is whether various thermodynamic quantities defined along single stochastic trajectories satisfy the large deviation principle [38,39]. The large deviations are concerned with the long-time fluctuation behavior of a stochastic process with small probability and it is closely related to the fluctuation theorem in the long-time limit. For Markovian systems, the large deviations for empirical measures, i.e., the number of times that each vertex of the graph is crossed per unit time, and for empirical flows, i.e., the number of times that each edge of the graph is traversed per unit time, have been extensively studied, while the large deviations for empirical cycle currents, i.e., the number of times that each cycle of the graph is formed per unit time, have received comparatively little attention. For cycle currents defined in the ST manner, the large deviations have been established since, in this case, the empirical cycle currents are exactly the empirical flows of chords [40,41]. For cycle currents defined in the LE manner, the explicit expression of the large deviation rate function is still unknown, even for systems with a simple topological structure.

In this paper, we make a comprehensive comparative study between cycle currents defined in the ST and LE manners, and clarify the connections and differences between them. The structure of this paper is organized as follows. In Sec. II, we recall the definitions of the two types of cycle currents and make a brief comparison between them. In Sec. III, we investigate the large deviations for the two types of cycle currents. We obtain the exact joint distribution and rate function for LE currents of a monocyclic Markovian system using the so-called cycle insertion method, and also obtain the exact rate function for ST currents of a general Markovian system. In Sec. IV, we state and compare various types of fluctuation theorems and symmetric relations satisfied by the two types of cycle currents. We clarify the ranges of applications of these fluctuation theorems and show that all the results for ST currents can be derived naturally from the relevant results for LE currents. We conclude in Sec. V.

II. MODEL AND TWO TYPES OF CYCLE CURRENTS

A. Model

Here we consider a thermodynamic system modeled by a discrete-time Markov chain $\xi = (\xi_n)_{n\geqslant 0}$ with state space $S = \{1, 2, \cdots, N\}$ and transition probability matrix $P = (p_{ij})_{i,j\in S}$, where p_{ij} denotes the transition probability from state i to state j. The transition diagram of the Markov chain ξ is a directed graph $G = (S, E)$, where the vertex set S is the state space and the edge set E contains all directed edges with positive transition probabilities (Fig. 1). In this paper, we use $\langle i, j\rangle$ to denote the edge from state i to state j. With this notation, the edge set E can be written more clearly as

$$E = \{\langle i, j\rangle \in S \times S : p_{ij} > 0\},$$

and we assume that $|E| = M$, where $|E|$ denotes the number of elements in E. Here we assume that the Markov chain ξ is irreducible, which means that G is a connected graph. Since the transition from a particular state to itself is allowed for a Markov chain, graph G may contain an edge from a state to itself, i.e., a self-loop (Fig. 1).

A special case occurs when the transition diagram G has a cyclic topology (except all self-loops), as illustrated in Fig. 1(c). Such systems will be referred to as monocyclic Markov chains in this paper. Specifically, the Markov chain ξ is called *monocyclic* if $p_{ij} = 0$ for any $|i - j| \geqslant 2$, where i and j are understood to be modulo N. In fact, monocyclic systems are of particular relevance in the biological context. Many crucial biochemical processes such as conformational changes of enzymes and ion channels [42,43], progression of cell cycle [44–46], phenotypic switching of cell types [47,48], phosphorylation-dephosphorylation cycle [49,50], and activation of promoters due to chromatin remodeling and transcription factor binding [51,52] can all be modeled as monocyclic Markov chains. In what follows, we mainly focus on monocyclic systems, while most of the results can be extended to general systems.

B. Cycle currents defined in the loop-erased manner

In this paper, we will investigate and compare two types of cycle currents. We first recall cycle currents defined in the LE manner [17,53]. A *circuit* of the Markov chain ξ is defined as a path $i_1 \to i_2 \to \cdots \to i_s \to i_1$ in graph G from a state to itself, where $i_1, i_2, \cdots, i_s$ are distinct states in S. Let $j_1 \to j_2 \to \cdots \to j_r \to j_1$ be another circuit. The above two circuits are said to be *equivalent* if $r = s$ and there exists an integer k such that

$$j_1 = i_{k+1}, j_2 = i_{k+2}, \cdots, j_r = i_{k+s},$$

where $k + 1, \cdots, k + s$ are understood to be modulo s. The equivalence class of the circuit $i_1 \to i_2 \to \cdots \to i_s \to i_1$ under the equivalence relation described above is called a *cycle* and is often denoted by $c = (i_1, i_2, \cdots, i_s)$. For example, (1,2,3), (2,3,1) and (3,1,2) represent the same cycle. The *reversed cycle* of $c = (i_1, i_2, \cdots, i_s)$ is defined as $c- = (i_1, i_s, \cdots, i_2)$. The set of all cycles is called the *cycle space* and is denoted by $\mathcal{C}$.

The trajectory of a Markov chain constantly forms various cycles. Intuitively, if we discard the cycles formed by ξ and

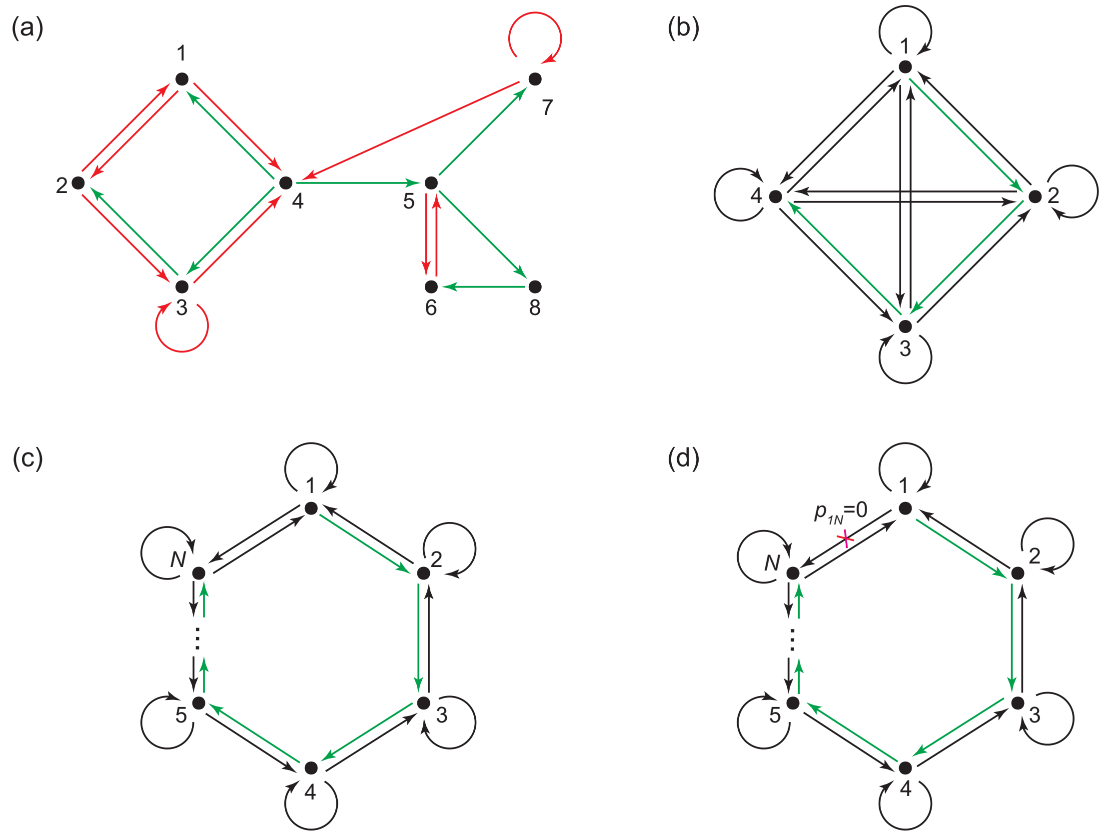

FIG. 1. Transition diagrams and the associated spanning trees for various Markov chains. (a) A Markov chain with a general transition diagram. The green arrows show the spanning tree T with root vertex 4, and the red arrows show all the chords of T. (b) A fully connected Markov chain with four states, where each state can transition to both itself and any other states. (c) A monocyclic Markov chain with N states. Each state can only transition to itself and its two neighbors. (d) A monocyclic Markov chain with N states. Here the transition from state 1 to state N is forbidden. In (b)–(d), the green arrows show the spanning tree T.

keep track of the remaining states in the trajectory, then we obtain a new Markov chain $\tilde{\xi} = (\tilde{\xi}_n)_{n\geqslant 0}$ called the *derived chain*. For example, if the trajectory of the original chain ξ is $\{1, 2, 3, 3, 2, 3, 4, 1, 4, \cdots\}$, then the corresponding trajectory of the derived chain $\tilde{\xi}$ and the cycles formed are shown in Table I.

More rigorously, a state of the derived chain $\tilde{\xi}$ is a finite sequence $i_1, i_2, \cdots, i_s$ of distinct states in S, denoted by $[i_1, i_2, \cdots, i_s]$. Suppose that $\tilde{\xi}_{n-1} = [i_1, i_2, \cdots, i_s]$ and $\xi_n = i_{s+1}$. If $i_{s+1} \notin \{i_1, i_2, \cdots, i_s\}$, then $\tilde{\xi}_n$ is defined as (see Table I for an illustration)

$$\tilde{\xi}_n = [i_1, i_2, \cdots, i_s, i_{s+1}].$$

On the other hand, if $i_{s+1} = i_r$ for some $1 \leqslant r \leqslant s$, then $\tilde{\xi}_n$ is defined as (see Table I for an illustration)

$$\tilde{\xi}_n = [i_1, i_2, \cdots, i_r].$$

In this case, we say that the Markov chain ξ forms cycle $c = (i_r, i_{r+1}, \cdots, i_s)$ at time n. Let N_n^c be the number of times that cycle c is formed up to time n. Then the *empirical (absolute) current* of cycle c up to time n is defined as

$$J_n^c = \frac{1}{n}N_n^c,$$

and the *empirical net current* of cycle c up to time n is defined as $\tilde{J}_n^c = J_n^c - J_n^{c-}$. Intuitively, J_n^c represents the number of times that cycle c is formed per unit time and $\tilde{J}_n^c$ represents the net number of times that cycle c is formed per unit time.

As $n \to \infty$, the empirical cycle current $J_n^c \to J^c$ and empirical net cycle current $\tilde{J}_n^c \to \tilde{J}^c$ will both converge with probability one. The limits J^c and $\tilde{J}^c$ are called the *current* and *net current* of cycle c, respectively. The explicit expressions of J_c and $\tilde{J}_c$ can be found in Ref. [17]. The well-known cycle

TABLE I. An example of the derived chain and the cycles formed.

n	0	1	2	3	4	5	6	7	8
ξ_n	1	2	3	3	2	3	4	1	4
$\tilde{\xi}_n$	[1]	[1,2]	[1,2,3]	[1,2,3]	[1,2]	[1,2,3]	[1,2,3,4]	[1]	[1,4]
Cycles formed				(3)	(2,3)			(1,2,3,4)	

current decomposition theorem [17] states that

$$\pi_i p_{ij} = \sum_{c \ni \langle i,j \rangle} J^c, \tag{1}$$

where π_i is the steady-state probability of state i and the sum on the right-hand side is taken over all cycles c which traverses edge $\langle i, j \rangle$ (the symbol $c \ni \langle i, j \rangle$ means that cycle c traverses edge $\langle i, j \rangle$). This shows that the probability flux between any pair of states can be decomposed as the sum of cycle currents.

C. Cycle currents defined in the spanning-tree manner

The current of a cycle can also be defined in the ST manner [8,53]. Let T be a directed subgraph of the transition diagram G, i.e., all the edges of T are also edges of G, and let $\overline{T}$ denote the undirected graph associated with T. Recall that T is called a *ST* (or maximal tree) of G if the following three conditions are satisfied [53]:

(a) T is a covering subgraph of G, i.e. T contains all the vertices of G;

(b) $\overline{T}$ is connected;

(c) $\overline{T}$ has no circuits, where a circuit of an undirected graph is defined as an undirected path from a vertex to itself.

In the following, we use T to represent both the ST itself and its edge set. The meaning should be clear from the context. In general, the choice of the ST is not unique, which means that a graph may have many different STs. It is easy to see that any ST T must contain all the vertices of G and must have $N - 1$ edges (see the green arrows in Fig. 1) [53].

A directed edge $l \notin T$ is called a *chord* of T [see the red arrows in Fig. 1(a)]. Since $|E| = M$ and $|T| = N - 1$, any ST T must have $M - N + 1$ chords. Since $\overline{T}$ is connected and has no circuits, if we add to T one of its chords l, then the resulting undirected subgraph $\overline{T \cup \{l\}}$ must have exactly one circuit. Let c_l be the cycle obtained from this circuit with the orientation being the same as chord l. For example, for the system illustrated in Fig. 1(a), if we add the chord $l = \langle 2, 1 \rangle$ to the ST T, then we obtain the cycle $c_l = (2, 1, 4, 3)$. The family of cycles $\mathcal{L} = \{c_l : l \notin T\}$ generated by the chords is referred to as the *fundamental set*. Since there is a one-to-one correspondence between the chord set and the fundamental set, the number of times that cycle c_l is formed is simply defined as the number of times that chord l is traversed. Along this line, the *empirical (absolute) current* of cycle c_l up to time n is defined as

$$Q_n^{c_l} = \frac{1}{n} \sum_{m=1}^{n} 1_{\{\langle \xi_{m-1}, \xi_m \rangle = l\}}.$$

Intuitively, $Q_n^{c_l}$ represents the number of times that chord l is traversed per unit time. Unlike the LE technique which can be used to define the currents of all cycles, the ST technique can only be used to define the currents of cycles in the fundamental set.

Similarly, we can define the empirical net current in the ST manner. The *empirical net current* of cycle c_l up to time n is defined as $\tilde{Q}_n^{c_l} = Q_n^{c_l} - Q_n^{c_l-}$. If c_l is composed of one or two states, then $c_l = c_l-$ and thus $\tilde{Q}_n^{c_l} = 0$. For any chord $l = \langle i, j \rangle$, if c_l is composed of three or more states and if c_l- is in the fundamental set, then $l- = \langle j, i \rangle$ must also be a chord and c_l- is exactly the cycle generated by the chord $l-$. As $n \to \infty$, the empirical cycle current $Q_n^{c_l} \to Q^{c_l}$ and empirical net cycle current $\tilde{Q}_n^{c_l} \to \tilde{Q}^{c_l}$ will both converge with probability one. The limits Q^{c_l} and $\tilde{Q}^{c_l}$ are called the *current* and *net current* of cycle c_l, respectively. For any chord $l = \langle i, j \rangle$, it follows from the ergodic theorem of Markov chains that $Q^{c_l} = \pi_i p_{ij}$.

We emphasize that most previous papers focused on net cycle currents defined in the LE [35] and ST [8,33] manners, and absolute cycle currents have received much less attention. Clearly, the net currents vanish for any one-state and two-state cycles. Hence, in previous papers [8,33,35], the net currents are only defined for cycles with three or more states. In this paper, we focus on both absolute and net currents. Here, following Refs. [17,53], we extend the definition slightly to include cycles with one or two states. This extension turns out to be useful, as can be seen in Sec. III below.

D. Comparisons between two types of cycle currents

Next we make a brief comparison between the two types of cycle currents. In what follows, cycle currents defined in the LE manner will be called *LE currents* and those defined in the ST manner will be called *ST currents*. We have seen that LE currents are defined for all cycles in the cycle space $\mathcal{C}$, while ST currents are only defined for cycles in the fundamental set $\mathcal{L}$. Hence, LE currents provide a more complete description of the cycle dynamics than ST currents. Moreover, since the ST is in general not unique, different choices of the ST correspond to different ST currents. Clearly, LE currents are independent of the choice of the ST.

A natural question is how much the fundamental set $\mathcal{L}$ is smaller than the cycle space $\mathcal{C}$. Since each chord corresponds to one and only one element in $\mathcal{L}$, we have $|\mathcal{L}| = M - N + 1$. It is difficult to provide a unified expression for $|\mathcal{C}|$. To gain deeper insights, we focus on two special cases. We first consider a Markov chain whose transition diagram is fully connected, i.e., $p_{ij} > 0$ for any $i, j \in S$, as illustrated in Fig. 1(b). In this case, the number of cycles with k states is given by $N(N-1)\cdots(N-k+1)/k$, and thus

$$|\mathcal{C}| = \sum_{k=1}^{N} \frac{N(N-1)\cdots(N-k+1)}{k}.$$

In particular, when $N = 4$, we have $|\mathcal{C}| = 24$ and the cycle space is given by

$$\begin{aligned}\mathcal{C} = \{&(1), (2), (3), (4), (1, 2), (1, 3), (1, 4), (2, 3), (2, 4),\\ &\times (3, 4), (1, 2, 3), (1, 2, 4), (1, 3, 2), (1, 3, 4), (1, 4, 2),\\ &\times (1, 4, 3), (2, 3, 4), (2, 4, 3), (1, 2, 3, 4), (1, 2, 4, 3),\\ &\times (1, 3, 2, 4), (1, 3, 4, 2), (1, 4, 2, 3), (1, 4, 3, 2)\}.\end{aligned}$$

If we choose the ST to be $T = 1 \to 2 \to 3 \to 4$, then $|\mathcal{L}| = 13$ and the fundamental set is given by

$$\begin{aligned}\mathcal{L} = \{&(1), (2), (3), (4), (1, 2), (2, 3), (3, 4)\\ &\times (1, 2, 3), (1, 3, 2), (2, 3, 4), (2, 4, 3), (1, 2, 3, 4,),\\ &\times (1, 4, 3, 2)\}.\end{aligned}$$

For a fully connected system, the number of ST currents is much smaller than the number of LE currents.

We next consider the monocyclic Markov chain illustrated in Fig. 1(c), where each state can only transition to itself and its two neighbors. In this case, we have $|\mathcal{C}| = 2N + 2$ and the cycle space is given by

$$\mathcal{C} = \{(1), \cdots, (N), (1,2), \cdots, (N-1,N), (N,1), \times (1,2,\cdots,N), (1,N,\cdots,2)\}. \quad (2)$$

The first N cycles are one-state cycles, i.e., self-loops, the middle N cycles are two-state cycles, and the last two cycles are N-state cycles. If we choose the ST to be $T = 1 \to 2 \to \cdots \to N$, then $|\mathcal{L}| = 2N + 1$ and the fundamental set is given by

$$\mathcal{L} = \{(1), \cdots, (N), (1,2), \cdots, (N-1,N), (1,2,\cdots,N), \times (1,N,\cdots,2)\}.$$

For a monocyclic system, there is only one cycle, i.e., cycle $(N,1)$, that is contained in $\mathcal{C}$ but is not contained in $\mathcal{L}$.

To further understand the relationship between the LE current J_n^c and the ST current $Q_n^{c_n}$, we use the convention of periodic boundary conditions, i.e., $\xi_0 = \xi_n$, which is a standard assumption in the literature [39]. With this assumption, for any chord l, it is easy to see that

$$Q_n^{c_l} = \sum_{c \ni l} J_n^c, \quad (3)$$

where the sum is taken over all cycles c that traverse chord l. Both sides of the equation represent the number of times that chord l is formed per unit time. This shows that ST currents can be represented as the sum of LE currents.

III. JOINT DISTRIBUTION AND LARGE DEVIATIONS FOR CYCLE CURRENTS

Previous studies about cycle currents mainly focused on the fluctuation relations, i.e., the symmetry relations satisfied by the probability distribution of cycle currents [35,36]. However, very little is known about the explicit expression of the probability distribution. Here we will address this problem and then use it to study the large deviations for cycle currents. In Sec. III A, we use methods in combinatorics and graph theory to compute the explicit expression of the joint probability distribution for LE currents. In Sec. III B, using the exact joint distribution and the Stirling formula, we investigate the large deviations for LE currents and give the explicit expression of the corresponding rate function. In Sec. III C, we study the large deviations for ST currents using the existing large deviation results for empirical flows.

A. Joint distribution for LE currents of monocyclic Markov chains

We first focus on the joint distribution for empirical LE currents $(J_n^c)_{c\in\mathcal{C}}$. In general, it is very difficult to obtain the explicit expression of the joint distribution for a general Markov chain. Here we focus on the monocyclic system illustrated in Fig. 1(c). All possible cycles formed by the system are listed in Eq. (2). Without loss of generality, we assume that the system starts from state 1. For each cycle $c = (i_1, i_2, \cdots, i_s)$, let $\gamma^c = p_{i_1 i_2} p_{i_2 i_3} \cdots p_{i_s i_1}$ denote the product of transition probabilities along this cycle. For any sequence of negative integers $k = (k^c)_{c\in\mathcal{C}}$ satisfying $\sum_{c\in\mathcal{C}} |c| k^c = n$, since we have assumed the periodic boundary condition, the joint distribution of empirical LE currents is given by

$$\mathbb{P}\left(J_n^c = \nu^c,\ \forall c \in \mathcal{C}\right) = \mathbb{P}\left(N_n^c = k^c,\ \forall c \in \mathcal{C}\right) = |G_n(k)| \prod_{c\in\mathcal{C}} (\gamma^c)^{k^c},$$

where $\nu^c = k^c/n$ is the frequency of occurrence of cycle c and $G_n(k)$ denotes the set of all possible trajectories up to time n so each cycle c is formed k^c times. Such trajectories will be called *allowable trajectories* in what follows. For convenience, we write k^c as k^i if $c = (i)$ is a one-state cycle, as $k^{i,i+1}$ if $c = (i, i+1)$ is a two-state cycle, as k^+ if $c = (1, 2, \cdots, N)$ is the clockwise N-state cycle, and as k^- if $c = (1, N, \cdots, 2)$ is the counterclockwise N-state cycle [Fig. 1(c)]. For example, for a three-state system, if the sequence $k = (k^c)_{c\in\mathcal{C}}$ is chosen as

$$k^3 = k^{12} = k^{23} = k^- = 1, \quad k^1 = k^2 = k^{13} = k^+ = 0, \quad (4)$$

then there are eight allowable trajectories up to time $n = 8$, and all of them are listed in Table II. Similarly, we write ν^c as ν^i, $\nu^{i,i+1}$, ν^+, and ν^- and write J^c as J^i, $J^{i,i+1}$, J^+, and J^-.

TABLE II. An example of allowable trajectories for a monocyclic system. All eight allowable trajectories for a three-state system up to time $n = 8$ so each one of the four cycles (3), (12), (23), and (1,3,2) are formed once, while the remaining four cycles (1), (2), (13), and (1,2,3) are not formed, i.e., $k^3 = k^{12} = k^{23} = k^- = 1$ and $k^1 = k^2 = k^{13} = k^+ = 0$.

m	0	1	2	3	4	5	6	7	8
ξ_m	1	3	3	2	3	2	1	2	1
ξ_m	1	3	2	3	3	2	1	2	1
ξ_m	1	3	3	2	1	2	3	2	1
ξ_m	1	3	2	1	2	3	3	2	1
ξ_m	1	2	3	3	2	1	3	2	1
ξ_m	1	2	3	2	1	3	3	2	1
ξ_m	1	2	1	3	3	2	3	2	1
ξ_m	1	2	1	3	2	3	3	2	1

We will next compute the number $|G_n(k)|$ of allowable trajectories. The basic idea is to insert all cycles into the trajectory in some appropriate order. The number of all possible insertions will then be the number of all allowable trajectories. The calculation is divided into the following three steps.

Step 1. Since we have assumed that the system starts from state 1, as the first step, we select all cycles containing the initial state 1, i.e., (1), (1,2), $(N, 1)$, $(1, 2, \cdots, N)$, $(1, N, \cdots, 2)$, and insert them into the trajectory. Since each cycle c is formed k^c times, the total number of possible insertions in step 1, i.e., the number of all permutations of these cycles, are given by

$$A_1 = \binom{k^1 + k^{12} + k^{N1} + k^+ + k^-}{k^1, k^{12}, k^{N1}, k^+, k^-} := \frac{(k^1 + k^{12} + k^{N1} + k^+ + k^-)!}{k^1!\, k^{12}!\, k^{N1}!\, k^+!\, k^-!}.$$

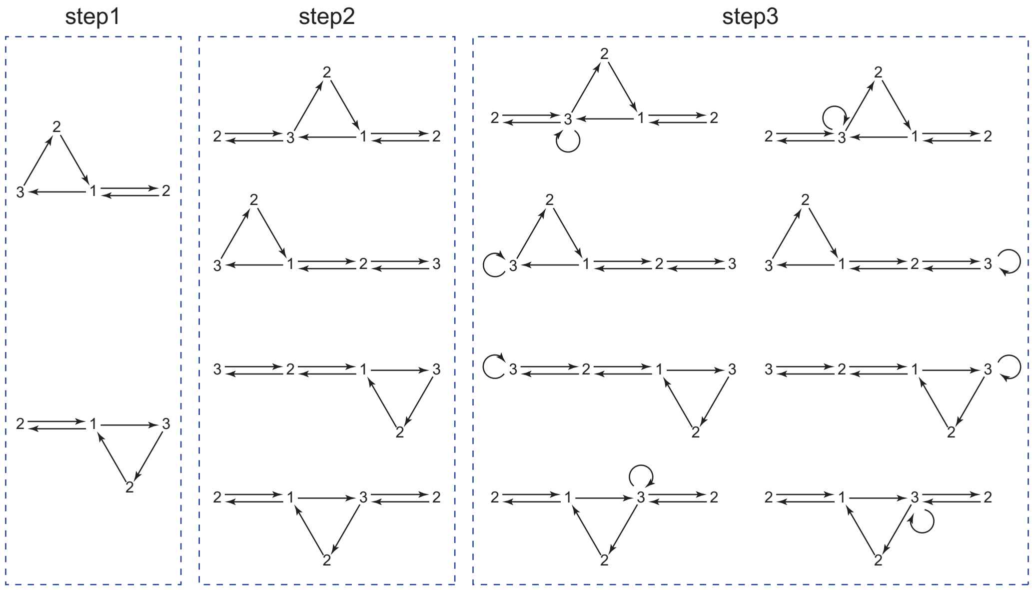

FIG. 2. Schematic of the cycle insertion method of constructing all allowable trajectories. Here we use the example given in Eqs. (4). The cycle insertion method is divided into three steps: First, we insert all cycles containing the initial state into the trajectory, next we insert all the remaining two-state cycles into the trajectory, and, finally, we insert all remaining one-state cycles into the trajectory. After the three-step cycle insertion, we find all eight allowable trajectories, which coincide exactly with those listed in Table II.

For the example given in Eqs. (4), all possible insertions in step 1 are shown in the left panel of Fig. 2.

Step 2. We next insert the remaining two-state cycles $(2,3),(3,4),\cdots,(N-1,N)$ into the trajectory. Note that when the system forms a two-state cycle $(i,i+1)$, it may be formed at state i or state $i+1$. For example, for the trajectory $\{1,3,2,3,\cdots\}$, when cycle (2,3) is formed, the derived chain becomes [1,3]. In this case, we say that the cycle is formed at state 3. On the contrary, for the trajectory $\{1,2,3,2,\cdots\}$, when cycle (2,3) is formed, the derived chain becomes [1,2]. In this case, we say that the cycle is formed at state 2.

For any two-state cycle $(i,i+1)$, let l^i and m^i denote the number of times that it is formed at state i and state $i+1$, respectively. Clearly, we have $l^i+m^i=k^{i,i+1}$. When l^i and m^i are fixed, the number of allowable trajectories can be computed as follows. First, we insert the l^2 cycle (2,3) at state 2. There are $k^{12}+k^+$ possible positions for the insertion, which correspond to state 2 in the cycles (1,2) and $(1,2,\cdots,N)$, which have been arranged in step 1. Note that these positions do not include state 2 in cycle $(1,N,\cdots,2)$. This is because if we insert cycle (2,3) here, then the cycle will be formed at state 3 rather than state 2. Hence the number of possible insertions is given by

$$\binom{k^{12}+k^++l^2-1}{l^2}. \tag{5}$$

Then we insert the l^i cycle $(i,i+1)$ at state i one by one for $3\leqslant i\leqslant N-1$. For each i, there are $l^{i-1}+k^+$ possible positions for the insertion, which correspond to state i in cycles $(i-1,i)$ and $(1,2,\cdots,N)$. The number of possible insertions is given by

$$\binom{l^{i-1}+k^++l^i-1}{l^i},\quad 3\leqslant i\leqslant N-1. \tag{6}$$

Thus far, we have inserted the l^i cycle $(i,i+1)$ at state i one by one for $2\leqslant i\leqslant N-1$. Combining Eqs. (5) and (6), the number of possible insertions is given by

$$\prod_{i=2}^{N-1}\binom{l^i+l^{i-1}+k^+-1}{l^i},$$

where $l^1:=k^{12}$.

Next we insert the m^i cycle $(i,i+1)$ at state $i+1$ one by one for $2\leqslant i\leqslant N-1$ in a similar way, and the number of possible insertions is given by

$$\prod_{i=2}^{N-1}\binom{m^i+m^{i+1}+k^--1}{m^i},$$

where $m^N:=k^{N1}$. Up till now, we have inserted all two-state cycles into the trajectory. Summing over all choices of l^i and m^i, the total number of possible insertions in step 2 is given by

$$A_2=\sum_{l^2+m^2=k^{23}}\cdots\sum_{l^{N-1}+m^{N-1}=k^{N-1,N}}\prod_{i=2}^{N-1}\binom{l^i+l^{i-1}+k^+-1}{l^i}\times\prod_{i=2}^{N-1}\binom{m^i+m^{i+1}+k^--1}{m^i}.$$

For the example given in Eqs. (4), all possible insertions in step 2 are shown in the middle panel of Fig. 2.

Step 3. We finally insert the remaining one-state cycles into the trajectory. Specifically, we insert cycle (i) into the trajectory one by one for $2\leqslant i\leqslant N$. For each i, there are $\sum_{c\ni i}k^c-k^i$ possible positions for the insertion, which correspond to state i in all the cycles except cycle (i). Hence the total number of possible insertions in step 3 is given by

$$A_3 = \prod_{i=2}^{N} \binom{\sum_{c\ni i} k^c - 1}{k^i}.$$

For the example given in Eqs. (4), all possible insertions in step 3 are shown in the right panel of Fig. 2.

Combining the above three steps, we finally obtain the number of allowable trajectories, which is given by

$$|G_n(k)| = A_1A_2A_3.$$

Hence, for a monocyclic system, the joint distribution of empirical LE currents can be computed exactly as

$$\mathbb{P}\left(J_n^c = \nu^c,\ \forall c\in\mathcal{C}\right) = A_1A_2A_3\prod_{c\in\mathcal{C}}(\gamma^c)^{k^c}. \tag{7}$$

We have seen that most previous papers [8,33,35] mainly focused on cycles with three or more states since the net currents for all one-state and two-state cycles must vanish. Here, we extend the definition slightly to include cycles with one and two states. This extension has the following two advantages: (i) in this paper, we not only focus on net cycle currents but also focus on absolute cycle currents; it is clear that the absolute currents for one-state and two-state cycles do not vanish and thus cannot be ignored; and (ii) only when all one-state and two-state cycles are taken into account, is it possible to recover all the allowable trajectories from empirical cycle currents using the three-step cycle insertion method; in this way, the joint distribution of empirical cycle currents has a simple closed-form expression.

B. Large deviations for LE currents of monocyclic Markov chains

The large deviations are concerned with the long-time fluctuation behavior of a stochastic process with small probability [38,39]. We next investigate the large deviations for empirical LE currents of a monocyclic Markov chain. Note that under the periodic boundary condition, the empirical LE currents $(J_n^c)_{c\in\mathcal{C}}$ must lie in the space

$$\mathcal{V} = \left\{(\nu^c)_{c\in\mathcal{C}}:\ \nu^c\geqslant 0,\ \sum_{c\in\mathcal{C}}|c|\nu^c = 1\right\},$$

where $|c|$ denotes the length of cycle c, i.e., the number of states contained in cycle c. Roughly speaking, $(J_n^c)_{c\in\mathcal{C}}$ are said to satisfy a large deviation principle with rate function $I_J:\mathcal{V}\to[0,\infty]$ if the joint distribution satisfies

$$\mathbb{P}\left(J_n^c = \nu^c,\ \forall c\in\mathcal{C}\right) \propto e^{-nI_J(\nu)},\quad n\to\infty \tag{8}$$

for any $\nu = (\nu^c)_{c\in\mathcal{C}}\in\mathcal{V}$. Clearly, the large deviation theory can capture the long-time fluctuation behavior of cycle currents. Next, we only present the main idea of the proof. The rigorous definition and proof of the large deviation principle can be found in Appendix A.

To obtain the explicit expression of the rate function I_J, we recall the Stirling formula

$$\log n! = n\log n - n + O(\log n) = h(n) - n + O(\log n),$$

where *log* represents the natural logarithm throughout the paper and $h(x) = x\log x$ for any $x\geqslant 0$. For convenience, we set $k_i = \sum_{c\ni i}k^c$ and $\nu_i = \sum_{c\ni i}\nu^c$. Note that the definitions of k_i and k^i are different. It then follows from the Stirling formula that

$$\begin{aligned}\log A_1 &= \log\frac{k_1!}{k^1!\,k^{12}!\,k^{N1}!\,k^+!\,k^-!}\\ &= h(k_1) - h(k^1) - h(k^{12}) - h(k^{N1})\\ &\quad - h(k^+) - h(k^-) + O(\log n)\\ &= n[h(\nu_1) - h(\nu^1) - h(\nu^{12}) - h(\nu^{N1}) - h(\nu^+)\\ &\quad - h(\nu^-)] + O(\log n).\end{aligned} \tag{9}$$

Similarly, we have

$$\begin{aligned}\log A_3 &= \log\prod_{i=2}^{N}\binom{k_i-1}{k^i} = \sum_{i=2}^{N}\log\frac{k_i!}{k^i!(k_i-k^i)!}\\ &= \sum_{i=2}^{N}[h(k_i) - h(k^i) - h(k_i-k^i)] + O(\log n)\\ &= \sum_{i=2}^{N}n[h(\nu_i) - h(\nu^i) - h(\nu_i-\nu^i)] + O(\log n).\end{aligned} \tag{10}$$

Finally, we estimate $\log A_2$. Let $D = \{(l^i,m^i)_{2\leqslant i\leqslant N-1}:\ l^i,m^i\in\mathbb{N},\ l^i+m^i = k^{i,i+1}\}$ denote the set of all possible choices of l^i and m^i. For any $L = (l^i,m^i)\in D$, let

$$B_L = \prod_{i=2}^{N-1}\binom{l^i+l^{i-1}+k^+-1}{l^i}\binom{m^i+m^{i+1}+k^--1}{m^i}$$

be the number of insertions in step 2 when l^i and m^i are fixed. It is clear that $|D|\leqslant n^{N-2}$. Thus, we have

$$\max_{L\in D}B_L \leqslant A_2 \leqslant n^{N-2}\max_{L\in D}B_L, \tag{11}$$

where we have used the fact that $A_2 = \sum_{L\in D}B_L$. Similarly to Eq. (10), we have

$$\begin{aligned}\log B_L &= \sum_{i=2}^{N-1}[h(l^i+l^{i-1}+k^+) - h(l^i) - h(l^{i-1}+k^+)]\\ &\quad + \sum_{i=2}^{N-1}[h(m^i+m^{i+1}+k^-) - h(m^i)\\ &\quad - h(m^{i+1}+k^-)] + O(\log n)\\ &= \sum_{i=2}^{N-1}n[h(x^i+x^{i-1}+\nu^+) - h(x^i) - h(x^{i-1}+\nu^+)]\\ &\quad + \sum_{i=2}^{N-1}n[h(y^i+y^{i+1}+\nu^-) - h(y^i) - h(y^{i+1}+\nu^-)]\\ &\quad + O(\log n),\end{aligned} \tag{12}$$

where $x^i = l^i/n$ and $y^i = m^i/n$. For any $\nu\in\mathcal{V}$, we introduce the space

$$V(\nu) = \{(x^i,y^i)_{2\leqslant i\leqslant N-1}:\ x^i,y^i\geqslant 0,\ x^i+y^i = \nu^{i,i+1}\},$$

and for any $X = (x^i, y^i) \in V(\nu)$, we define the function

$$F_\nu(X) = \sum_{i=2}^{N-1} [h(x^i) + h(x^{i-1} + \nu^+) - h(x^i + x^{i-1} + \nu^+)] + \sum_{i=2}^{N-1} [h(y^i) + h(y^{i+1} + \nu^-) - h(y^i + y^{i+1} + \nu^-)], \tag{13}$$

where $x^1 = \nu^{12}$ and $y^N = \nu^{N1}$. It then follows from Eq. (11) that

$$\log A_2 = \max_{L\in D} \log B_L + O(\log n) = n \sup_{X\in V(\nu)} F_\nu(X) + O(\log n). \tag{14}$$

Combining Eqs. (7) and (8), we obtain

$$I_J(\nu) = -\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(J_n^c = \nu^c,\ \forall c \in \mathcal{C}\right) = -\lim_{n\to\infty} \frac{1}{n}\left[\log A_1 + \log A_2 + \log A_3 + \sum_{c\in\mathcal{C}} k^c \log \gamma^c\right].$$

It then follows from Eqs. (9), (10), and (14) that

$$I_J(\nu) = [h(\nu^{12}) + h(\nu^{N1}) + h(\nu^+) + h(\nu^-) - h(\nu^{12} + \nu^{N1} + \nu^+ + \nu^-)] + \inf_{X\in V(\nu)} F_\nu(X) + \sum_{i\in S}[h(\nu_i - \nu^i) + h(\nu^i) - h(\nu_i)] - \sum_{c\in\mathcal{C}} \nu^c \log \gamma^c, \tag{15}$$

where $h(x) = x\log x$ and $\nu_i = \sum_{c\ni i} \nu^c$. This gives the expression of the rate function I_J for empirical LE currents. We emphasize that the explicit expressions of the joint probability distribution and the rate function for empirical LE currents, i.e., Eqs. (7) and (15), are only applicable to monocyclic Markov chains and cannot be applied to general systems. For a general Markov chain, the process of cycle formation is much more intricate and cannot be computed using the three-step cycle insertion method.

Note that in Eq. (15), it is difficult to compute the term $\inf_{X\in V(\nu)} F_\nu(X)$. A more explicit expression of this term can be obtained using the Lagrange multiplier method. In Appendix B, we have proved that

$$\inf_{X\in V(\nu)} F_\nu(X) = F_\nu(x^i, y^i),$$

where $(x^i, y^i)_{2\leqslant i\leqslant N-1}$ is any solution (such solution must exist but may not be unique) of the following set of algebraic equations:

$$\frac{x^i}{x^{i-1} + x^i + \nu^+} \cdot \frac{x^i + \nu^+}{x^i + x^{i+1} + \nu^+} = \frac{y^i + \nu^-}{y^{i-1} + y^i + \nu^-} \cdot \frac{y^i}{y^i + y^{i+1} + \nu^-},$$
$$x^i + y^i = \nu^{i,i+1}, \tag{16}$$

with $x^1 = \nu^{12}$, $x^N = 0$, $y^1 = 0$, and $y^N = \nu^{N1}$.

Thus far, we have assumed that the system starts from state 1. A natural question is whether the rate function will change when the system starts from other initial distributions. In fact, we can prove that the rate function is independent of the choice of the initial distribution. Note that this is a highly nontrivial result because in the expression Eq. (15), the status of state 1 and the status of other states are not equal. The proof is rather complicated and is put in Appendix C.

For a general monocyclic system, the expression Eq. (15) of the rate function is very complicated. This expression can be greatly simplified in two special cases: (i) the case where the system has only three states (any three-state system must be monocyclic) and (ii) the case where the transition from state 1 to state N is forbidden [see Fig. 1(d) for an illustration]. For a three-state system, the rate function reduces to (see Appendix D for the proof)

$$I_J(\nu) = \sum_{i\in S}\left[\nu^i \log\left(\frac{\nu^i/\nu_i}{J^i/J_i}\right) + (\nu_i - \nu^i)\log\left(\frac{(\nu_i - \nu^i)/\nu_i}{(J_i - J^i)/J_i}\right)\right] + \sum_{c\in\mathcal{C},|c|\neq 1} \nu^c \log\left(\frac{\nu^c/\tilde{\nu}}{J^c/\tilde{J}}\right), \tag{17}$$

where

$$\tilde{\nu} = \sum_{c\in\mathcal{C},|c|\neq 1} \nu^c = \nu^{12} + \nu^{13} + \nu^{23} + \nu^+ + \nu^-,$$
$$\tilde{J} = \sum_{c\in\mathcal{C},|c|\neq 1} J^c = J^{12} + J^{13} + J^{23} + J^+ + J^-.$$

For an N-state monocyclic system with the transition from state 1 to state N being forbidden [Fig. 1(d)], the rate function reduces to (see Appendix D for the proof)

$$I_J(\nu) = \sum_{i\in S}\left[\nu^i \log\left(\frac{\nu^i/\nu_i}{J^i/J_i}\right) + \nu^{i,i+1}\log\left(\frac{\nu^{i,i+1}/\nu_i}{J^{i,i+1}/J_i}\right) + (\nu^{i-1,i} + \nu^+)\log\left(\frac{(\nu^{i-1,i} + \nu^+)/\nu_i}{(J^{i-1,i} + J^+)/J_i}\right)\right]. \tag{18}$$

Note that the expressions of the rate function in the two special cases are much simpler and more symmetric than the general expression given in Eq. (15). Clearly, both expressions have a symmetric form with respect to each state and thus are independent of the choice of the initial distribution. It is well-known that the empirical flows of a Markov chain, i.e., the number of times that each edge is traversed per unit time, also satisfy a large deviation principle and the associated rate function has the form of relative entropy (see Sec. III C for details). Interestingly, we find that the rate functions given in Eqs. (17) and (18) also have a functional form similar to relative entropy.

The large deviations for empirical LE currents $(J_n^c)_{c\in\mathcal{C}}$ can be directly applied to establish the large deviations for empirical net LE currents $(\tilde{J}_n^c)_{c\in\mathcal{C}}$. Since the empirical net LE currents vanish for any one-state and two-state cycles and since $\tilde{J}_n^+ = -\tilde{J}_n^-$ for the two N-state cycles $(1, 2, \cdots, N)$ and $(1, N, \cdots, 2)$, we only need to focus on the empirical net currents $\tilde{J}_n^+$ of cycle $(1, 2, \cdots, N)$. By the contraction principle, we have

$$\begin{aligned}\mathbb{P}(\tilde{J}_n^+ = x) &= \mathbb{P}(J_n^+ - J_n^- = x)\\ &= \sum_{\nu^+-\nu^-=x} \mathbb{P}\big(J_n^c = \nu^c, \forall c \in \mathcal{C}\big)\\ &\propto \sum_{\nu^+-\nu^-=x} e^{-nI_J(\nu)}, \quad n\to\infty. \end{aligned} \tag{19}$$

This shows that the empirical net LE current $\tilde{J}_n^+$ satisfies a large deviation principle with rate function

$$I_{\tilde{J}}(x) = \inf_{\{\nu\in\mathcal{V}:\, \nu^+-\nu^-=x\}} I_J(\nu). \tag{20}$$

C. Large deviations for ST currents of general Markov chains

We next focus on the large deviations for empirical ST currents of a general Markov chain. In fact, the large deviations for empirical net ST currents have been investigated and the symmetry of the rate function has been obtained in Ref. [40]. Here we focus on the large deviations for empirical (absolute) ST currents. To this end, we first recall the large deviations for empirical flows [39].

Recall that the empirical flow of edge $\langle i, j\rangle$ up to time n is defined as

$$R_n(i,j) = \frac{1}{n}\sum_{m=1}^n 1_{\{\xi_{m-1}=i,\xi_m=j\}}.$$

Intuitively, $R_n(i, j)$ represents the number of times that edge $\langle i, j\rangle$ is traversed per unit time. Note that under the periodic boundary condition, the empirical flows $(R_n(i, j))_{\langle i,j\rangle\in E}$ must lie in the space

$$\mathcal{M} = \Bigg\{(R_n(i,j))_{\langle i,j\rangle\in E} : R_n(i,j)\geqslant 0, \sum_{i,j\in S} R(i,j) = 1, \sum_{j\in S} R(i,j) = \sum_{j\in S} R(j,i)\Bigg\}.$$

It is well-known that the empirical flows $(R_n(i, j))_{\langle i,j\rangle\in E}$ satisfy the following large deviation principle:

$$\mathbb{P}(R_n(i,j) = R(i,j),\ \forall\langle i,j\rangle\in E) \propto e^{-nI_{\rm flow}(R)}, \quad n\to\infty,$$

where the rate function $I_{\rm flow} : \mathcal{M}\to[0,\infty]$ is given by

$$I_{\rm flow}(R) = \sum_{\langle i,j\rangle\in E} R(i,j)\log\frac{R(i,j)}{R(i)p_{ij}},$$

with $R(i) = \sum_{j\in S} R(i, j)$. Clearly, the rate function for empirical flows has the form of relative entropy. For any chord l of a fixed ST T, let H^{c_l} be a function on E defined by

$$H^{c_l}(i,j) = \begin{cases} 1, & \text{if (i) } \langle i,j\rangle\in T \text{ and } \langle i,j\rangle\in c_l \text{ or (ii) } \langle i,j\rangle = l\\ -1, & \text{if } \langle i,j\rangle\in T,\ \langle i,j\rangle\notin c_l, \text{ and } \langle j,i\rangle\in c_l\\ 0, & \text{otherwise.}\end{cases} \tag{21}$$

In fact, the empirical flow $R_n(i, j)$ can be represented as the weighted sum of $H^{c_l}(i, j)$ with the weights being all empirical ST currents [53], i.e.,

$$R_n(i,j) = \sum_{c_l\in\mathcal{L}} Q_n^{c_l} H^{c_l}(i,j), \quad \langle i,j\rangle\in E.$$

It was further proved in Ref. [53] that this representation is unique. In other words, if $R_n = \sum_{c_l\in\mathcal{L}} \mu^{c_l} H^{c_l}$ for some coefficients μ^{c_l}, then we must have $\mu^{c_l} = Q_n^{c_l}$ for any $c_l\in\mathcal{L}$. It then follows from the uniqueness of the above representation that

$$\begin{aligned}&\mathbb{P}\big(Q_n^{c_l} = \mu^{c_l},\ \forall c_l\in\mathcal{L}\big)\\ &= \mathbb{P}\Bigg(R_n(i,j) = \sum_{c_l\in\mathcal{L}} \mu^{c_l}H^{c_l}(i,j),\ \forall\langle i,j\rangle\in E\Bigg)\\ &\propto e^{-nI_{\rm flow}\left(\sum_{c_l\in\mathcal{L}}\mu^{c_l}H^{c_l}\right)}, \quad n\to\infty.\end{aligned}$$

This shows that the empirical ST currents $(Q_n^{c_l})_{c_l\in\mathcal{L}}$ satisfy a large deviation principle with rate function

$$I_Q(\mu) = I_{\rm flow}\Bigg(\sum_{c_l\in\mathcal{L}}\mu^{c_l}H^{c_l}\Bigg). \tag{22}$$

Thus far, we have obtained the explicit expressions of the rate function for empirical LE currents of a monocyclic system and the rate function for empirical ST currents of a general system. A natural question is what the relationship is between the two rate functions. To see this, recall that ST currents can be represented by LE currents as $Q_n^{c_l} = \sum_{c\ni l} J_n^c$. It thus follows from the contraction principle that

$$\begin{aligned}\mathbb{P}\big(Q_n^{c_l} = \mu^{c_l},\ \forall l\in\mathcal{L}\big) &= \mathbb{P}\Bigg(\sum_{c\ni l} J_n^c = \mu^{c_l},\ \forall l\in\mathcal{L}\Bigg)\\ &= \sum_{\sum_{c\ni l}\nu^c=\mu^{c_l}} \mathbb{P}\big(J_n^c = \nu^c,\ \forall c\in\mathcal{C}\big)\\ &\propto \sum_{\sum_{c\ni l}\nu^c=\mu^{c_l}} e^{-nI_J(\nu)}, \quad n\to\infty.\end{aligned}$$

This shows that the rate functions for empirical LE and ST currents are connected by

$$I_Q(\mu) = \inf_{\{\nu\in\mathcal{V}:\, \sum_{c\ni l}\nu^c=\mu^{c_l}\}} I_J(\nu).$$

It is straightforward to prove that the rate function I_Q given above coincides with the one given in Eq. (22) for monocyclic systems.

The large deviations for empirical ST currents $(Q_n^{c_l})_{c_l\in\mathcal{L}}$ can also be used to establish the large deviations for empirical net ST currents $(\tilde{Q}_n^{c_l})_{c_l\in\mathcal{L}}$. Since the empirical net ST currents vanish for all one-state and two-state cycles, we only need to focus on cycles with three or more states. Let $c_{l_1}, c_{l_2}, \cdots, c_{l_s}$ be all cycles with three or more states in the fundamental set so any two of them are not reversed cycles of each other. By the contraction principle, the empirical net ST currents $(\tilde{Q}^{c_{l_i}})_{1\leqslant i\leqslant s}$ of these cycles satisfy a large deviation principle with rate function

$$I_{\tilde{Q}}(x) = \inf_{\{\mu\in\mathcal{M}:\, \mu^{c_{l_i}}-\mu^{c_{l_i}^-}=x_i,\ \forall 1\leqslant i\leqslant s\}} I_Q(\mu). \tag{23}$$

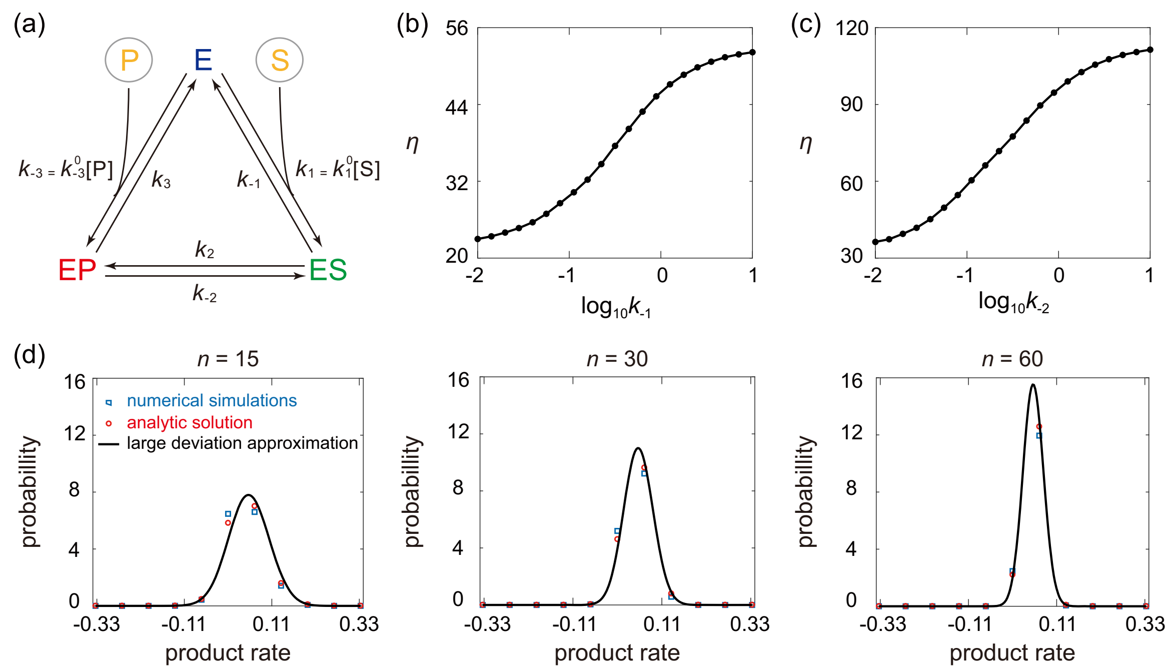

FIG. 3. Three-step mechanism of a reversible enzyme reaction. (a) Kinetic scheme of a three-step reversible enzyme reaction. Here k_1^0 and k_{-3}^0 are second-order rate constants since they are the rate constants of the second-order reactions $E + S \to ES$ and $E + P \to EP$, respectively, while $k_1 = k_1^0[S]$ and $k_{-3} = k_{-3}^0[P]$ are pseudo-first-order rate constants since they are the rate constants of the effective first-order reactions $E \to ES$ and $E \to EP$, respectively (here we have assumed that the system is open and $[S]$ and $[P]$ are sustained by an external agent). From the perspective of a single enzyme molecule, the reaction is unimolecular and cyclic. (b) Noise η in the product rate as k_{-1} varies. The parameters are chosen as $\tau = 0.01, n = 15, k_1^0 = 2k_{-1}, k_2 = 1, k_{-2} = 1, k_3 = 1, k_{-3}^0 = 0.1, [P] = 1$, and $[S]$ is tuned so $\langle \tilde{J}_n^+ \rangle$ remains invariant. (c) Noise η in the product rate as k_{-2} varies. The parameters are chosen as $\tau = 0.01, n = 15, k_1^0 = 1.2, k_{-1} = 0.6, k_2 = k_{-2}, k_3 = 1, k_{-3}^0 = 0.1$, $[P] = 1$, and $[S]$ is tuned so$\langle \tilde{J}_n^+ \rangle$ remains invariant. (d) Distribution of the product rate as time n increases. The blue squares are the ones obtained using stochastic simulations, the red circles are the ones obtained using the exact joint distribution Eq. (7), and the black curves are the ones obtained using the exact rate function Eq. (17) and large deviation approximation Eqs. (8). The parameters are chosen as $\tau = 0.1$, $k_1^0 = 4, k_{-1} = 2, k_2 = 5, k_{-2} = 1, k_3 = 6, k_{-3}^0 = 0.1$, and $[S] = [P] = 1$.

D. Applications in single-molecule enzyme kinetics

As an application of our theoretical results, we consider the following three-step mechanism of a reversible enzyme reaction [10,54],

$$E + S \underset{k_{-1}}{\overset{k_1^0}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EP \underset{k_{-3}^0}{\overset{k_3}{\rightleftharpoons}} E + P,$$

where E is an enzyme turning the substrate S into the product P. Here k_1^0 and k_{-3}^0 are second-order rate constants since they are the rate constants of second-order reactions, while k_{-1}, $k_{\pm 2}$, and k_3 are first-order rate constants since they correspond to first-order reactions. If there is only one enzyme molecule, then it may convert stochastically among three conformal states: the free enzyme E, the enzyme-substrate complex ES, and the enzyme-product complex EP. For simplicity, we assume that the enzyme reaction is in an open system with the concentrations of S and P sustained by an external agent [49]. Then from the enzyme perspective, the kinetics is stochastic and cyclic with pseudo-first-order rate constants $k_1 = k_1^0[S]$ and $k_{-3} = k_{-3}^0[P]$, where $[S]$ and $[P]$ are the sustained concentrations of S and P, respectively [Fig. 3(a)]. Note that the time variable of the enzyme reaction is continuous. However, in experiments, we are only able to observe the system at multiple discrete time points. If we record the conformal state of the enzyme molecule at a series of time points with interval τ, then the system can be modeled as a three-state discrete-time Markov chain, which coincides with the model studied in this paper. Let $Q = (q_{ij})$ be the transition rate matrix of the continuous-time system shown in Fig. 3(a). Then the transition probability matrix of the discrete-time system is given by $P = (p_{ij}) = e^{\tau Q}$ [55]. The discrete-time system serves as a good approximation of the continuous-time system when the interval τ is small.

Note that for the cyclic kinetics illustrated in Fig. 3(a), a substrate molecule S is converted into a product molecule P whenever the clockwise cycle $C^+ = (E, ES, EP)$ is formed, and a product molecule P is converted into a substrate molecule S whenever the counterclockwise cycle $C^- = (E, EP, ES)$ is formed. Thus the rate of product formation, also called *product rate*, of the enzyme reaction, i.e., the net conversion of S into P per unit time, is exactly the net LE current $\tilde{J}_n^+ = J_n^+ - J_n^-$. Previous studies [10] mainly focus on the long-time mean product rate

$$\lim_{n\to\infty} \tilde{J}_n^+ = \tilde{J}^+ = \frac{\gamma^+ - \gamma^-}{C},$$

where $\gamma^+ = p_{12}p_{23}p_{31}$, $\gamma^- = p_{13}p_{32}p_{21}$, and

$$C = \sum_{i=1}^{3}[(1 - p_{i-1,i-1})(1 - p_{i+1,i+1}) - p_{i-1,i+1}p_{i+1,i-1}].$$

The analytical results derived in previous sections allow us to investigate the finite-time fluctuation behavior of the product rate $\tilde{J}_n^+$. In experiments, the size of the fluctuations, also called *noise*, in the product rate is often measured by the coefficient of variation $\eta = \sigma/\mu$, where $\mu = \langle \tilde{J}_n^+ \rangle$ is the mean and σ is the standard deviation [56,57]. Note that we have obtained the exact joint distribution of empirical LE currents in Sec. III. Using the joint distribution, it is easy to calculate all moments, including the mean and standard deviation, of the product rate $\tilde{J}_n^+$.

In Figs. 3(b) and 3(c), we illustrate noise η as a function of the rate constants k_{-1} and k_{-2}. Here k_{-1} and k_{-2} are varied while keeping k_1^0/k_{-1} and k_2/k_{-2} as constant, and the substrate concentration $[S]$ is tuned so the mean product rate μ remains invariant (examining protein noise while fixing the protein mean is a common strategy in molecular biology experiments [58]). From Fig. 3(b), we see that noise in the product rate becomes larger as k_1^0 and k_{-1} increase (while keeping their ratio as constant). Note that when k_1^0 and k_{-1} are both large, the reaction $E + S \rightleftharpoons ES$ will reach rapid pre-equilibrium and this is widely known as rapid equilibrium assumption in enzyme kinetics [50]. Our results show that rapid equilibrium between enzyme states E and ES leads to large fluctuations in the product rate. Similarly, from Fig. 3(c), we find that noise in the product rate also becomes larger as k_2 and k_{-2} increase (while keeping their ratio as constant). Note that when k_2 and k_{-2} are both large, the two enzyme states ES and EP will reach rapid pre-equilibrium and thus can be combined into a single state [59,60]. In this case, the three-step enzyme reaction reduces to the classical two-step Michaelis-Menten enzyme kinetics

$$E + S \rightleftharpoons ES \rightleftharpoons E + P.$$

This implies that compared to the two-step Michaelis-Menten kinetics, the three-step kinetics results in smaller fluctuations in the product rate.

While the exact joint distribution for LE currents derived in Sec. III A can be used to study the fluctuations in the product rate, it is computationally very slow because we need to calculate a large number of factorials and combinatorial numbers [see Eq. (7)], especially when time n and the number of states N are large. Fortunately, the large deviations for LE currents studied in Sec. III B can be used to provide a much more efficient computational method of the joint distribution. Specifically, we only need to compute the rate function $I_J(\nu)$ using Eq. (17) and then apply Eqs. (8) to construct an approximation of the joint distribution. In Fig. 3(d), we compare the distribution of the production rate $\tilde{J}_n^+$ obtained by using stochastic simulations (blue squares), the analytical solution (red circles), and the large deviation approximation (black curves). As expected, the analytical solution coincides perfectly with stochastic simulations. Interestingly, we find that the approximate distribution obtained based on the large deviation theory is in good agreement with the analytical solution when $n \geqslant 15$ and they become practically indistinguishable when $n \geqslant 30$. According to our simulations, when $n = 30$, compared with the analytical solution, the large deviation approximation can save the computational time by over 99%. This suggests that the large deviation principle studied in this paper is very useful because it enables a fast exploration of large swaths of parameter space.

IV. FLUCTUATION THEOREMS FOR CYCLE CURRENTS

Next we investigate the fluctuation relations satisfied by the two types of cycle currents. In Sec. IV A, using trajectory reversal method, we obtain a symmetric relation for LE currents of a monocyclic system that is even stronger than the classical transient and integral fluctuation theorems. In Sec. IV B, we generalize the fluctuation relations to a general system and reveal their connection with the second law of thermodynamics. In Sec. IV C, we explore the fluctuation relations for ST currents of a general system and compare them with the fluctuation relations for LE currents.

A. Fluctuation theorems for LE currents of monocyclic Markov chains

An important question is whether empirical cycle currents satisfy various fluctuation theorems. In fact, the transient fluctuation theorem for net LE currents has been investigated in Refs. [35,61]. Here we will prove a symmetric relation for a monocyclic system that is even stronger than the transient fluctuation theorem. For convenience, we write the two N-state cycles of a monocyclic system as $C^+ = (1, 2, \cdots, N)$ and $C^- = (1, N, \cdots, 2)$. Let N_n^+ and N_n^- denote the number of times that cycles C^+ and C^- are formed up to time n, respectively. The strong symmetric relation for LE currents is given by

$$\begin{aligned} &k^+\mathbb{P}\left(N_n^+ = k^+, N_n^- = k^- - 1, N_n^c = k^c,\ \forall c \neq C^+, C^-\right) \\ &\quad = \left(\frac{\gamma^+}{\gamma^-}\right)k^-\mathbb{P}\left(N_n^+ = k^+ - 1, N_n^- = k^-, N_n^c = k^c,\right. \\ &\quad\quad \left.\forall c \neq C^+, C^-\right), \end{aligned} \tag{24}$$

where $\gamma^+ = p_{12}p_{23}\cdots p_{N1}$ and $\gamma^- = p_{21}p_{32}\cdots p_{N1}$ are the product of transition probabilities along cycles C^+ and C^-, respectively. In fact, a similar equality has been obtained recently for another type of cycle currents defined in the sequence matching manner [20]. We next give the proof of Eq. (24) for LE currents. Under the periodic boundary condition, it follows from Eq. (7) that

$$\begin{aligned} &\mathbb{P}\left(N_n^+ = k^+, N_n^- = k^- - 1, N_n^c = k^c,\ \forall c \neq C^+, C^-\right) \\ &\quad = (\gamma^+)^{k^+}(\gamma^-)^{k^- - 1} \\ &\quad\quad \times \prod_{c\neq C^+,C^-} (\gamma^c)^{k^c}|G_n(k^+, k^- - 1, (k^c)_{c\neq C^+,C^-})|, \end{aligned}$$

where $G_n(k^+, k^- - 1, (k^c)_{c\neq C^+,C^-})$ is the collection of all possible trajectories up to time n so cycle C^+ is formed k^+ times, cycle C^- is formed $k^- - 1$ times, and any other cycle $c \neq C^+, C^-$ is formed k^c times. For simplicity of notation, we rewrite the above equation as

$$\begin{aligned} &\mathbb{P}(N_n^+ = k^+, N_n^- = k^- - 1, \cdots) \\ &\quad = (\gamma^+)^{k^+}(\gamma^-)^{k^- - 1} \prod_{c\neq C^+,C^-} (\gamma^c)^{k^c}|G_n(k^+, k^- - 1, \cdots)|. \end{aligned}$$

Similarly, replacing k^+ by $k^+ - 1$ and replacing $k^- - 1$ by k^- in the above equation, we obtain

$$\begin{aligned}&\mathbb{P}(N_n^+ = k^+ - 1, N_n^- = k^-, \cdots)\\&\quad = (\gamma^+)^{k^+-1}(\gamma^-)^{k^-}\prod_{c\neq C^+,C^-}(\gamma^c)^{k^c}|G_n(k^+ - 1, k^-, \cdots)|.\end{aligned}$$

Hence, to prove Eq. (24), we only need to show that

$$k^+|G_n(k^+, k^- - 1, \cdots)| = k^-|G_n(k^+ - 1, k^-, \cdots)|. \quad (25)$$

For any trajectory $\{\xi_0, \xi_1, \cdots, \xi_n\}$ lying in $G_n(k^+, k^- - 1, \cdots)$, since cycle C^+ is formed k^+ times, there are k^+ beginning times (the times that C^+ begins to form) and k^+ ending times (the times that C^+ has been formed) for this cycle. Let T_i^{begin} and T_i^{end} denote the ith beginning and ending times for cycle C^+, respectively. For example, for the trajectory given in Table I, the first beginning time for cycle $c = (1, 2, 3, 4)$ is $n = 0$ and the first ending time is $n = 7$. If we reverse the trajectory $\{\xi_0, \xi_1, \cdots, \xi_n\}$ between T_i^{begin} and T_i^{end}, then we obtain a new trajectory $\{\tilde{\xi}_0, \tilde{\xi}_1, \cdots, \tilde{\xi}_n\}$, which is given by

$$\tilde{\xi}_m = \begin{cases}\xi_{T_i^{\text{begin}}+T_i^{\text{end}}-m}, & \text{if } T_i^{\text{begin}} \leqslant m \leqslant T_i^{\text{end}}\\ \xi_m, & \text{otherwise.}\end{cases}$$

Clearly, the reversed trajectory must lie in $G_n(k^+ - 1, k^-, \cdots)$. Since cycle C^+ is formed k^+ times, there are $k^+|G_n(k^+, k^- - 1, \cdots)|$ possible reversed trajectories. Among these reversed trajectories, k^- trajectories are exactly the same and are counted repetitively. For example, if $C^+ = (1, 2, 3)$ and $C^- = (1, 3, 2)$, then the trajectories $\{1, 2, 3, 1, 2, 3, 1, 3, 2\}$ and $\{1, 2, 3, 1, 3, 2, 1, 2, 3\}$ in $G_8(2, 1, \cdots)$ can both be reversed to the trajectory $\{1, 2, 3, 1, 3, 2, 1, 3, 2\}$ in $G_8(1, 2, \cdots)$, and thus are counted twice. As a result, the number of possible trajectories in $G_n(k^+ - 1, k^-, \cdots)$ is given by

$$|G_n(k^+ - 1, k^-, \cdots)| = \frac{k^+}{k^-}|G_n(k^+, k^- - 1, \cdots)|,$$

which is exactly Eq. (25). Thus we have proved the strong symmetric relation given by Eq. (24). Applying the symmetric relation $|k^+ - k^-|$ times, we obtain the transient fluctuation theorem for LE currents:

$$\begin{aligned}&\mathbb{P}(N_n^+ = k^+, N_n^- = k^-, \cdots)\\&\quad = \mathbb{P}(N_n^+ = k^-, N_n^- = k^+, \cdots)\left(\frac{\gamma^+}{\gamma^-}\right)^{k^+-k^-}. \quad (26)\end{aligned}$$

Thus far, we have proved the symmetric relation Eq. (24) and transient fluctuation theorem Eq. (26) under the periodic boundary condition. Without the periodic boundary condition, these two equalities are also valid for monocyclic systems; the proof is similar and thus is omitted.

The transient fluctuation theorem can be used to prove the other two types of fluctuation theorems. To see this, recall that the moment-generating function of empirical LE currents is defined as

$$g_n(\lambda^+, \lambda^-, \cdots) = \left\langle e^{\lambda^+N_n^+ + \lambda^-N_n^- + \sum_{c\neq C^+,C^-}\lambda^cN_n^c}\right\rangle,$$

where $\langle A\rangle$ denotes the mean of A. Then the following Kurchan-Lebowitz-Spohn-type fluctuation theorem holds:

$$\begin{aligned}&g_n(\lambda^+, \lambda^-, \cdots)\\&= \sum_k e^{\sum_{c\in C}\lambda^ck^c}\mathbb{P}(N^+ = k^+, N^- = k^-, \cdots)\\&= \sum_k e^{\sum_{c\in C}\lambda^ck^c}\mathbb{P}(N^+ = k^-, N^- = k^+, \cdots)\left(\frac{\gamma^+}{\gamma^-}\right)^{k^+-k^-}\\&= \sum_k e^{\cdots+\left(\lambda^+-\log\frac{\gamma^+}{\gamma^-}\right)k^+ + \left(\lambda^- - \log\frac{\gamma^-}{\gamma^+}\right)k^-}\\&\quad\times\mathbb{P}(N^+ = k^-, N^- = k^+, \cdots)\\&= \left\langle e^{\left(\lambda^- - \log\frac{\gamma^+}{\gamma^-}\right)N_n^+ + \left(\lambda^+ + \log\frac{\gamma^+}{\gamma^-}\right)N_n^- + \cdots}\right\rangle\\&= g_n\left(\lambda^- - \log\frac{\gamma^+}{\gamma^-}, \lambda^+ + \log\frac{\gamma^+}{\gamma^-}, \cdots\right), \quad (27)\end{aligned}$$

where $\log(\gamma^+/\gamma^-)$ is the affinity of cycle C^+ [2]. We next consider the long-time limit behavior of a monocyclic system. As $n \to \infty$, it is easy to see that

$$\begin{aligned}e^{-nI_J(\nu^+,\nu^-,\cdots)} &\propto \mathbb{P}(J_n^+ = \nu^+, J_n^- = \nu^-, \cdots)\\&= \mathbb{P}(J_n^+ = \nu^-, J_n^- = \nu^+, \cdots)\left(\frac{\gamma^+}{\gamma^-}\right)^{n(\nu^+-\nu^-)}\\&\propto e^{-n[I_J(\nu^-,\nu^+,\cdots)-(\log\frac{\gamma^+}{\gamma^-})(\nu^+-\nu^-)]}.\end{aligned}$$

This yields the Gallavotti-Cohen-type fluctuation theorem:

$$I_J(\nu^+, \nu^-, \cdots) = I_J(\nu^-, \nu^+, \cdots) - \left(\log\frac{\gamma^+}{\gamma^-}\right)(\nu^+ - \nu^-). \quad (28)$$

Similarly, we can also obtain the fluctuation theorems for net LE currents. For a monocyclic system, we only need to focus on the empirical net LE current $\tilde{J}_n^+$ of cycle C^+. Let $\tilde{g}_n(\lambda) = \langle e^{\lambda n\tilde{J}_n^+}\rangle$ be the moment-generating function of $\tilde{J}_n^+$ and let $I_{\tilde{J}}(x)$ be the rate function of $\tilde{J}_n^+$ given in Eq. (20). The various types of fluctuation theorems for net LE currents follow directly from Eqs. (26)–(28) and are summarized as follows (these identities were first obtained in Ref. [61]).

(1) Transient fluctuation theorem:

$$\frac{\mathbb{P}(\tilde{J}_n^+ = x)}{\mathbb{P}(\tilde{J}_n^+ = -x)} = \left(\frac{\gamma^+}{\gamma^-}\right)^{nx}.$$

(2) Kurchan-Lebowitz-Spohn-type fluctuation theorem:

$$\tilde{g}_n(\lambda) = \tilde{g}_n\left(-\left(\lambda + \log\frac{\gamma^+}{\gamma^-}\right)\right).$$

(3) Integral fluctuation theorem: Taking $\lambda = -\log(\gamma^+/\gamma^-)$ in the above equation yields

$$\langle e^{\lambda n\tilde{J}_n^+}\rangle = 1.$$

(4) Gallavotti-Cohen-type fluctuation theorem:

$$I_{\tilde{J}}(x) = I_{\tilde{J}}(-x) - \left(\log\frac{\gamma^+}{\gamma^-}\right)x.$$

B. Fluctuation theorems for LE currents of general Markov chains

We have seen that various symmetric relations and fluctuation theorems hold for LE currents of a monocyclic system. A natural question is whether these results can be extended to a general Markov chain. Before stating the results, we recall the definition of similar cycles [36]. Let $c_1 = (i_1, i_2, \cdots, i_s)$ and $c_2 = (j_1, j_2, \cdots, j_r)$ be two cycles. Then c_1 and c_2 are called *similar* if $s = r$ and $\{i_1, i_2, \cdots, i_s\} = \{j_1, j_2, \cdots, j_r\}$. In other words, two cycles are similar if they pass through the same set of states. For example, the following six cycles:

$$(1,2,3,4),\ (1,2,4,3),\ (1,3,2,4),\ (1,3,4,2),\ (1,4,2,3),\ (1,4,3,2)$$

are similar. Note that any cycle C and its reversed cycle $C-$ must be similar.

We first focus on empirical LE currents $(J_n^c)_{c\in\mathcal{C}}$, where $J_n^c = N_n^c/n$. For a general Markov chain, if cycles c_1 and c_2 are similar, then the following symmetric relation holds:

$$\frac{k^{c_1}\mathbb{P}\left(N_n^{c_1} = k^{c_1}, N_n^{c_2} = k^{c_2}-1, N_n^c = k^c,\ \forall c \neq c_1, c_2\right)}{k^{c_2}\mathbb{P}\left(N_n^{c_1} = k^{c_1}-1, N_n^{c_2} = k^{c_2}, N_n^c = k^c,\ \forall c \neq c_1, c_2\right)} = \frac{\gamma^{c_1}}{\gamma^{c_2}}. \tag{29}$$

If we choose c_1 and c_2 to be some cycle C^+ and its revered cycle C^-, then this equality reduces to

$$\frac{k^{+}\mathbb{P}\left(N_n^{+} = k^{+}, N_n^{-} = k^{-}-1, N_n^c = k^c,\ \forall c \neq C^+, C^-\right)}{k^{-}\mathbb{P}\left(N_n^{+} = k^{+}-1, N_n^{-} = k^{-}, N_n^c = k^c,\ \forall c \neq C^+, C^-\right)} = \frac{\gamma^{+}}{\gamma^{-}}.$$

This can be viewed as a generalization of Eq. (24) in the monocyclic case. Applying Eq. (29) repeatedly gives the following transient fluctuation theorem for LE currents:

$$\frac{\mathbb{P}\left(N_n^{c_1} = k^{c_1}, N_n^{c_2} = k^{c_2}, N_n^c = k^c,\ \forall c \neq c_1, c_2\right)}{\mathbb{P}\left(N_n^{c_1} = k^{c_2}, N_n^{c_2} = k^{c_1}, N_n^c = k^c,\ \forall c \neq c_1, c_2\right)} = \left(\frac{\gamma^{c_1}}{\gamma^{c_2}}\right)^{k^{c_1}-k^{c_2}}. \tag{30}$$

This shows that if cycles c_1 and c_2 are similar, then the joint distribution of empirical LE currents satisfies a symmetric relation under the exchange of k^{c_1} and k^{c_2}. Actually, the proof of Eq. (30) has been given in Ref. [36] under the restrictions that all cycles under consideration pass through a common state $i \in S$ and the Markov chain also starts from state i. Fortunately, this technical assumption can be removed and the result holds generally (manuscript in preparation).

We next consider empirical net LE currents $(\tilde{J}_n^c)_{c\in\mathcal{C}}$. Let $c_1, c_2, \cdots, c_r$ be all cycles with three or more states in the cycle space so any two of them are not reversed cycles of each other (the empirical net LE currents for one-state and two-state cycles vanish and do not need to be considered). It then follows from Eq. (30) that

$$\begin{aligned}
&\mathbb{P}\left(\tilde{J}_n^{c_1} = x_1, \tilde{J}_n^{c_m} = x_m,\ \forall 2 \leqslant m \leqslant r\right)\\
&= \mathbb{P}\left(N_n^{c_1} - N_n^{c_1-} = nx_1, N_n^{c_m} - N_n^{c_m-} = nx_m,\ \forall 2 \leqslant m \leqslant r\right)\\
&= \sum_{k^{c_i}-k^{c_i-}=nx_i,\ \forall 1\leqslant i\leqslant r} \mathbb{P}\left(N_n^{c_1} = k^{c_1}, N_n^{c_1-} = k^{c_1-}, N_n^{c_m} = k^{c_m}, N_n^{c_m-} = k^{c_m-},\ \forall 2 \leqslant m \leqslant r\right)\\
&= \sum_{k^{c_i}-k^{c_i-}=nx_i,\ \forall 1\leqslant i\leqslant r} \mathbb{P}\left(N_n^{c_1} = k^{c_1-}, N_n^{c_1-} = k^{c_1}, N_n^{c_m} = k^{c_m}, N_n^{c_m-} = k^{c_m-},\ \forall 2 \leqslant m \leqslant r\right)\left(\frac{\gamma^{c_1}}{\gamma^{c_1-}}\right)^{nx_1}\\
&= \mathbb{P}\left(N_n^{c_1} - N_n^{c_1-} = -nx_1, N_n^{c_m} - N_n^{c_m-} = nx_m,\ \forall 2 \leqslant m \leqslant r\right)e^{nx_1\log\frac{\gamma^{c_1}}{\gamma^{c_1-}}}\\
&= \mathbb{P}\left(\tilde{J}_n^{c_1} = -x_1, \tilde{J}_n^{c_m} = x_m,\ \forall 2 \leqslant m \leqslant r\right)e^{nx_1\log\frac{\gamma^{c_1}}{\gamma^{c_1-}}}.
\end{aligned}$$

Hence we have obtained the following transient fluctuation theorems for net LE currents:

$$\frac{\mathbb{P}\left(\tilde{J}_n^{c_1} = x_1, \tilde{J}_n^{c_m} = x_m,\ \forall 2 \leqslant m \leqslant r\right)}{\mathbb{P}\left(\tilde{J}_n^{c_1} = -x_1, \tilde{J}_n^{c_m} = x_m,\ \forall 2 \leqslant m \leqslant r\right)} = e^{nx_1\log\frac{\gamma^{c_1}}{\gamma^{c_1-}}}. \tag{31}$$

This shows that the joint distribution of empirical net LE currents satisfies a symmetric relation when any x_i is replaced by $-x_i$. In fact, this result which was first found in Ref. [61] for a monocyclic system and further generalized in Ref. [35] to a general system, while the proof is not totally rigorous. If we change x_i to $-x_i$ one by one for $1 \leqslant i \leqslant r$ in the above equation, then we obtain

$$\frac{\mathbb{P}\left(\tilde{J}_n^{c_1} = x_1, \tilde{J}_n^{c_2} = x_2, \cdots, \tilde{J}_n^{c_r} = x_r\right)}{\mathbb{P}\left(\tilde{J}_n^{c_1} = -x_1, \tilde{J}_n^{c_2} = -x_2, \cdots, \tilde{J}_n^{c_r} = -x_r\right)} = e^{n\sum_{i=1}^r x_i\log\frac{\gamma^{c_i}}{\gamma^{c_i-}}}. \tag{32}$$

Note that Eq. (32) is much weaker than Eq. (31). In what follows, we term Eq. (31) the *strong form* and term Eq. (32) the *weak form* of the transient fluctuation theorem.

Other types of fluctuation theorems for absolute and net LE currents can be easily derived from the transient fluctuation theorem and are summarized as follows. Here we only focus on the strong form of various fluctuation theorems; the weak form can be obtained similarly. Let $g_n(\lambda) = \langle e^{n\sum_{c\in C}\lambda_i J_n^{c_i}}\rangle$ and $\tilde{g}_n(\lambda) = \langle e^{n\sum_{i=1}^r \lambda_i \tilde{J}_n^{c_i}}\rangle$ be the moment generating functions of $(J_n^{c_i})_{c_i\in\mathcal{C}}$ and $(\tilde{J}_n^{c_i})_{1\leqslant i\leqslant r}$, respectively. Moreover, let $I_J(x)$ and $I_{\tilde{J}}(x)$ be the rate functions of $(J_n^{c_i})_{c_i\in\mathcal{C}}$ and $(\tilde{J}_n^{c_i})_{1\leqslant i\leqslant r}$, respectively.

(1) Kurchan-Lebowitz-Spohn-type fluctuation theorem: if cycles c_1 and c_2 are similar, then

$$g_n(\lambda_1,\lambda_2,\cdots) = g_n\left(\lambda_2 - \log\frac{\gamma^{c_1}}{\gamma^{c_2}}, \lambda_1 + \log\frac{\gamma^{c_1}}{\gamma^{c_2}},\cdots\right),$$

$$\tilde{g}_n(\lambda_1,\cdots) = \tilde{g}_n\left(-\left(\lambda_1 + \log\frac{\gamma^{c_1}}{\gamma^{c_1-}}\right),\cdots\right).$$

(2) Integral fluctuation theorem: for any subset $\{c_1,c_2,\cdots,c_t\}\subset\{c_1,c_2,\cdots,c_r\}$, we have

$$\left\langle e^{-n\sum_{i=1}^t \tilde{J}_n^{c_i}\log\frac{\gamma^{c_i}}{\gamma^{c_i-}}}\right\rangle = 1. \quad (33)$$

(3) Gallavotti-Cohen-type fluctuation theorem: If cycles c_1 and c_2 are similar, then

$$I_J(x_1,x_2,\cdots) = I_J(x_2,x_1,\cdots) - \left(\log\frac{\gamma^{c_1}}{\gamma^{c_2}}\right)(x_1-x_2),$$

$$I_{\tilde{J}}(x_1,\cdots) = I_{\tilde{J}}(-x_1,\cdots) - \left(\log\frac{\gamma^{c_1}}{\gamma^{c_1-}}\right)x_1.$$

The fluctuation theorems for net LE currents have important physical implications. To see this, recall that the total entropy production of a Markovian system along a single trajectory $\{\xi_0,\xi_1,\cdots,\xi_n\}$ is given by [29]

$$\begin{aligned} S_n^{\text{tot}} &= \log\frac{\mu_0(\xi_0)p_{\xi_0\xi_1}p_{\xi_1\xi_2}\cdots p_{\xi_{n-1}\xi_n}}{\mu_n(\xi_n)p_{\xi_n\xi_{n-1}}p_{\xi_{n-1}\xi_{n-2}}\cdots p_{\xi_1\xi_0}} \\ &= \log\frac{\mu_0(\xi_0)}{\mu_n(\xi_n)} + \sum_{k=0}^{n-1}\log\frac{p_{\xi_k\xi_{k+1}}}{p_{\xi_{k+1}\xi_k}}, \end{aligned}$$

where $\mu_0 = (\mu_0(i))_{i\in S}$ is the distribution of ξ_0 and $\mu_n = (\mu_n(i))_{i\in S}$ is the distribution of ξ_n. Under the periodic boundary condition, it is clear that $\mu_0(\xi_0) = \mu_n(\xi_n)$. Moreover, we have

$$p_{\xi_0\xi_1}p_{\xi_1\xi_2}\cdots p_{\xi_{n-1}\xi_n} = \prod_{c\in\mathcal{C}}(\gamma^c)^{N_n^c},$$

$$p_{\xi_n\xi_{n-1}}p_{\xi_{n-1}\xi_{n-2}}\cdots p_{\xi_1\xi_0} = \prod_{c\in\mathcal{C}}(\gamma^c)^{N_n^{c-}}.$$

Combining the above two equations, we obtain

$$\begin{aligned} S_n^{\text{tot}} &= n\sum_{c\in\mathcal{C}}\tilde{J}_n^c\log\gamma^c = \frac{n}{2}\sum_{c\in\mathcal{C}}\tilde{J}_n^c\log\frac{\gamma^c}{\gamma^{c-}} \\ &= n\sum_{i=1}^r\tilde{J}_n^{c_i}\log\frac{\gamma^{c_i}}{\gamma^{c_i-}}, \end{aligned} \quad (34)$$

where we have used the fact that $\tilde{J}_n^{c-} = -\tilde{J}_n^c$ in the second identity. This shows that the total entropy production can be decomposed as the weighted sum of net LE currents with the weights being all cycle affinities, and the quantity $n\tilde{J}_n^c\log(\gamma^c/\gamma^c)$ can be understood as the entropy product along cycle c. It is well-known, as the total entropy production of any Markovian system satisfies the integral fluctuation theorem $\langle e^{-S_n^{\text{tot}}}\rangle = 1$ [29], which implies the classical second law of thermodynamics $\langle S_n^{\text{tot}}\rangle \geqslant 0$. Our results indicate that the integral fluctuation theorem not only holds for the total entropy production but also holds for the entropy production along any finite number of cycles $c_1,c_2,\cdots,c_t$ [see Eq. (33)]. In particular, for any cycle c, we have

$$\left\langle e^{-n\tilde{J}_n^c\log\frac{\gamma^c}{\gamma^{c-}}}\right\rangle = 1.$$

This is much stronger than the classical result for the total entropy production. Moreover, applying Jensen's inequality to the integral fluctuation theorem Eq. (33), we find

$$\left\langle\sum_{i=1}^t\tilde{J}_n^{c_i}\log\frac{\gamma^{c_i}}{\gamma^{c_i-}}\right\rangle \geqslant 0, \quad (35)$$

where $c_1,c_2,\cdots,c_t$ are any finite number of cycles. In particular, for any cycle c, we have

$$\left\langle\tilde{J}_n^c\log\frac{\gamma^c}{\gamma^{c-}}\right\rangle \geqslant 0.$$

This provides a much refined version of the second law of thermodynamics, which shows that the entropy production along any finite number of cycles has a nonnegative mean. This reveals the hidden refined structure behind the underlying system.

C. Fluctuation theorems for ST currents of general Markov chains

We have seen that both absolute and net LE currents satisfy various fluctuation theorems. A natural question is whether similar relations also hold for absolute and net ST currents. In fact, (absolute) ST currents do not satisfy any form of fluctuation theorems, even for monocyclic systems. To see this, consider a fully connected three-state system and let $T = 1\to 2\to 3$ be the ST. Then the fundamental set is given by

$$\mathcal{L} = \{(1),(2),(3),(1,2),(2,3),(1,2,3),(1,3,2)\}.$$

It then follows from Eq. (22) that the rate function for empirical ST currents is given by

$$I_Q(\mu) = \sum_{\langle i,j\rangle\in E}R^\mu(i,j)\log\frac{R^\mu(i,j)}{R^\mu(i)p_{ij}},$$

where $R^\mu(i,j) = \sum_{c_l\in\mathcal{L}}\mu^{c_l}H^{c_l}(i,j)$ and $R^\mu(i) = \sum_{j\in S}R^\mu(i,j)$. For simplicity of notation, let $\mu^+ = \mu^{(1,2,3)}$ and let $\mu^- = \mu^{(1,3,2)}$. In Fig. 4(a), we illustrate the difference between $I_Q(\mu^+,\mu^-,\cdots)$ and $I_Q(\mu^-,\mu^+,\cdots) - \log(\gamma^+/\gamma^-)(\mu^+-\mu^-)$ as a function of μ^+ and μ^- under

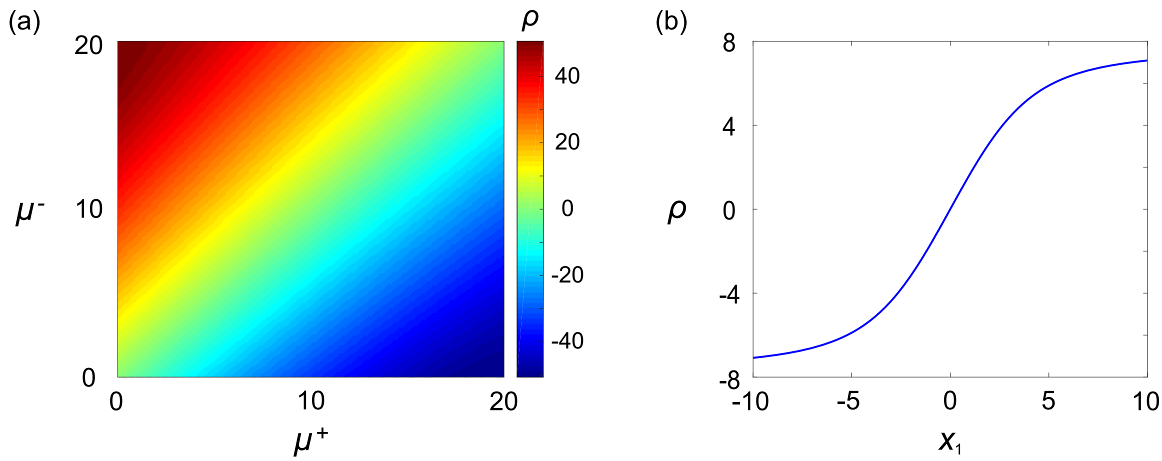

FIG. 4. Some fluctuation theorems may be broken for absolute and net ST currents. (a) Heat plot of $\rho = I_Q(\mu^+, \mu^-, \cdots) - I_Q(\mu^-, \mu^+, \cdots) + \log(\gamma^+/\gamma^-)(\mu^+ - \mu^-)$ as a function of μ^+ and μ^- for a three-state system. The fact that $\rho \not\equiv 0$ shows that the Gallavotti-Cohen-type fluctuation theorem is broken for absolute ST currents. The parameters are chosen as $\mu^1 = 15$, $\mu^2 = 20$, $\mu^3 = 3$, $\mu^{12} = 21$, $\mu^{23} = 37$, $p_{11} = 0.28$, $p_{12} = 0.22$, $p_{13} = 0.5$, $p_{21} = 0.1$, $p_{22} = 0.6$, $p_{23} = 0.3$, $p_{31} = 0.3$, $p_{32} = 0.3$, $p_{33} = 0.4$. (b) Change of $\rho = I_{\tilde{Q}}(x_1, x_2, x_3) - I_{\tilde{Q}}(-x_1, x_2, x_3) + \log(\gamma^{c_1}/\gamma^{c_1-})x_1$ as a function of x_1 for a four-state system. The fact that $\rho \not\equiv 0$ shows that the strong form of the Gallavotti-Cohen-type fluctuation theorem is broken for net ST currents. The parameters are chosen as $x_2 = 2$, $x_3 = 3$, $p_{11} = 0.1$, $p_{12} = 0.2$, $p_{13} = 0.3$, $p_{14} = 0.4$, $p_{21} = 0.5$, $p_{22} = 0.15$, $p_{23} = 0.15$, $p_{24} = 0.2$, $p_{31} = 0.1$, $p_{32} = 0.4$, $p_{33} = 0.25$, $p_{34} = 0.25$, $p_{41} = 0.2$, $p_{42} = 0.2$, $p_{43} = 0.3$, $p_{44} = 0.3$.

a set of appropriately chosen parameters. It is clear that the difference is nonzero and thus we have

$$I_Q(\mu^+, \mu^-, \cdots) \neq I_Q(\mu^-, \mu^+, \cdots) - \left(\log \frac{\gamma^+}{\gamma^-}\right)(\mu^+ - \mu^-), \tag{36}$$

which means that the Gallavotti-Cohen-type fluctuation theorem is broken. Other types of fluctuation theorems must also be broken since the Gallavotti-Cohen-type fluctuation theorem is the weakest among all fluctuation theorems.

While various fluctuation theorems fail for ST currents, they may hold for net ST currents [33]. To see this, note that for a monocyclic system, we only need to consider the empirical net current $\tilde{Q}_n^+$ of cycle C^+. Suppose that the ST is chosen as $T = 1 \to 2 \to \cdots \to N$. With the periodic boundary condition, it follows from Eq. (3) that $Q_n^+ = J_n^+ + J_n^{(N,1)}$ and $Q_n^- = J_n^- + J_n^{(N,1)}$. These two equations imply that $\tilde{Q}_n^+ = \tilde{J}_n^+$, and thus the fluctuation theorems for net ST currents naturally follow from those for net LE currents. Without the periodic boundary condition, the Gallavotti-Cohen-type fluctuation theorem still holds since it reflects the long-time behavior of the system and assuming the periodic boundary condition or not will not influence the large deviation rate function, while the other three types of fluctuation theorems are all broken. It has been shown in Ref. [34] that all four types of fluctuation theorems are satisfied for a modified version of net ST currents.

The above results can be extended to a general system. Let $c_{l_1}, c_{l_2}, \cdots, c_{l_s}$ be all cycles with three or more states in the fundamental set so any two of them are not reversed cycles of each other (the empirical net ST currents for one-state and two-state cycles vanish and do not need to be considered). In Ref. [33], the authors have proved the following weak form of the Gallavotti-Cohen-type fluctuation theorem for net ST currents:

$$I_{\tilde{Q}}(x_1, x_2, \cdots, x_s) = I_{\tilde{Q}}(-x_1, -x_2, \cdots, -x_s) - \sum_{i=1}^{s} x_i \log \frac{\gamma^{c_{l_i}}}{\gamma^{c_{l_i}-}}. \tag{37}$$

This shows that the joint distribution of empirical net ST currents satisfies a symmetric relation when all x_i are replaced by $-x_i$. In fact, the above equality can be obtained directly from the fluctuation theorems for net LE currents. For any cycle $c_l \in \mathcal{L}$ with three or more states, under the periodic boundary condition, it follows from Eq. (3) that

$$\tilde{Q}_n^{c_l} = \sum_{c \ni l} J_n^c - \sum_{c \ni l-} J_n^c = \sum_{c \ni l} J_n^c - \sum_{c \ni l} J_n^{c-} = \sum_{c \ni l} \tilde{J}_n^c. \tag{38}$$

This indicates that empirical net ST currents can be decomposed as the sum of empirical net LE currents. It then follows from Eq. (32) that (see Appendix E for the proof)

$$\frac{\mathbb{P}\left(\tilde{Q}_n^{c_{l_1}} = x_1, \cdots, \tilde{Q}_n^{c_{l_s}} = x_s\right)}{\mathbb{P}\left(\tilde{Q}_n^{c_{l_1}} = -x_1, \cdots, \tilde{Q}_n^{c_{l_s}} = -x_s\right)} = e^{n \sum_{i=1}^{s} x_i \log \frac{\gamma^{c_{l_i}}}{\gamma^{c_{l_i}-}}}. \tag{39}$$

This shows that net ST currents satisfy the weak form of the transient fluctuation theorem under the periodic boundary condition. The weak form Eq. (37) of the Gallavotti-Cohen-type fluctuation theorem holds generally since assuming the periodic boundary condition or not will not influence the large deviation rate function.

In contrast to net LE currents, net ST currents do not satisfy the strong form of fluctuation theorems; this fact has been found in previous papers [62–65]. To give a counterexample, we consider a fully connected four-state system illustrated in Fig. 1(b). Suppose that the ST is chosen as $T = 1 \to 2 \to 3 \to 4$. In this case, we only need to consider the net ST currents of the three cycles $c_1 = (1, 2, 3)$, $c_2 = (2, 3, 4)$, and

$c_3 = (1, 2, 3, 4)$, since other cycles in the fundamental set are either their reversed cycles or cycles with one or two states. Recall that the rate function of empirical net ST currents $(\tilde{Q}_n^{c_1}, \tilde{Q}_n^{c_2}, \tilde{Q}_n^{c_3})$ is given by

$$I_{\tilde{Q}}(x) = \inf_{\{\mu\in\mathcal{M}:\mu^{c_i}-\mu^{c_i-}=x_i,\ \forall 1\leqslant i\leqslant 3\}} I_Q(\mu). \tag{40}$$

In Fig. 4(b), we illustrate the difference between $I_{\tilde{Q}}(x_1, x_2, x_3)$ and $I_{\tilde{Q}}(-x_1, x_2, x_3) - \log(\gamma^{c_1}/\gamma^{c_1-})\tilde{\mu}^{c_1}$ as a function of x_1 under a set of appropriately chosen parameters. It is clear that the difference is nonzero and thus

$$I_{\tilde{Q}}(x_1, x_2, x_3) \neq I_{\tilde{Q}}(-x_1, x_2, x_3) - \left(\log\frac{\gamma^{c_1}}{\gamma^{c_1-}}\right)\tilde{\mu}^{c_1}. \tag{41}$$

Hence the strong form of the Gallavotti-Cohen-type fluctuation theorem fails for net ST currents.

We next discuss the connection between ST currents and entropy production. Similarly to Eq. (34), under the periodic boundary condition, the total entropy product along a single trajectory can be also decomposed as the weighted sum of net ST currents [8], i.e.,

$$S_n^{\text{tot}} = n\sum_{i=1}^{s}\tilde{Q}_n^{c_{l_i}}\log\frac{\gamma^{c_{l_i}}}{\gamma^{c_{l_i}-}}. \tag{42}$$

Hence within the ST framework, the quantity $n\tilde{Q}_n^{c_l}\log(\gamma^{c_l}/\gamma^{c_l-})$ can be understood as the entropy production *along fundamental cycle* c_l. Note that this is totally different from the quantity $\tilde{J}_n^{c_l}\log(\gamma^{c_l}/\gamma^{c_l-})$ investigated in Sec. IV B. We have seen that within the LE framework, the entropy production along any finite number of cycles satisfies both the strong form of integral fluctuation theorem Eq. (33) and the refined version of the second law of thermodynamics Eq. (35). Since the strong form of fluctuation theorems fails for net ST currents, the entropy production *along any fundamental cycle* does not satisfy the refined version of the second law of thermodynamics. In other words, it may occur that

$$\left\langle\tilde{Q}_n^{c_l}\log\frac{\gamma^{c_l}}{\gamma^{c_l-}}\right\rangle < 0$$

for some fundamental cycle c_l.

The reason why the strong form of fluctuation theorems and the refined version of the second law of thermodynamics are broken for net ST currents can be explained as follows. From Eq. (38), it is clear that the net ST current $\tilde{Q}_n^{c_l}$ of fundamental cycle c_l can be decomposed as the sum of the net LE currents $\tilde{J}_n^c$ of all cycles c that traverse chord l, i.e., $\tilde{Q}_n^{c_l} = \sum_{c\ni l}\tilde{J}_n^c$. Note that these cycles c that traverse chord l have different affinities, which may not be equal to the affinity of fundamental cycle c_l. Hence, even if $\langle\tilde{J}_n^c\log(\gamma^c/\gamma^{c-})\rangle \geqslant 0$ for all cycles c, we cannot conclude that $\langle\tilde{Q}_n^{c_l}\log(\gamma^{c_l}/\gamma^{c_l-})\rangle \geqslant 0$. The weak form of fluctuation theorems holds for net ST currents since it is essentially the fluctuation theorems for the total entropy production [see Eq. (42)].

In summary, we have seen that LE currents have much better properties than ST currents; the former satisfies a much more refined version of the second law of thermodynamics while the latter does not. This demonstrates the advantage of LE currents in dealing with complex thermodynamic systems far from equilibrium (for simple monocyclic systems, the net LE and ST currents are the same).

V. CONCLUSIONS AND DISCUSSION

In this paper, we make a comparative study of the large deviations and fluctuation theorems for empirical cycle currents of a Markov chain defined in the LE and ST manners. LE currents are defined for all cycles in the cycle space, while ST currents are only defined for cycles in the fundamental set generated by the chords of an arbitrarily chosen ST. The fundamental set may be much smaller than the cycle space for a general system. However, for a system with a cyclic topology, there is at most one cycle that is contained in the cycle space but is missing in the fundamental set. LE currents provide a more complete and detailed description for the cycle dynamics than ST currents. Under the periodic boundary condition, the ST current of any cycle can be represented by the weighted sum of LE currents.

Furthermore, we establish the large deviation principle and provide the explicit expression of the associated rate function for empirical LE currents of a monocyclic Markov chain. The proof is based on deriving the joint distribution of empirical LE currents of all cycles in closed form. When computing the joint distribution, we propose the method of three-step cycle insertion: (i) the first step is to insert all cycles that pass through the initial state into the trajectory, (ii) the second step is to insert all two-state cycles that do not contain the initial state into the trajectory, and (iii) the third step is to insert all one-state cycles that do not contain the initial state into the trajectory. In addition, the rate function is proved to be independent of the initial distribution of the system. The analytical expression of the rate function is complicated for a general monocyclic system. However, it can be greatly simplified for a three-state system and for a monocyclic system with a certain transition between adjacent states being forbidden. Whether the rate function of LE currents can be computed in a nonmonocyclic system remains an open question. Following the method proposed in Ref. [40], which only focused on empirical net ST currents, we also give the exact rate function for empirical (absolute) ST currents of a general system. The relationship between the rate functions of empirical LE and ST currents is clarified.

The analytical results are then applied to investigate the fluctuations in the product rate for a three-step reversible enzyme reaction, which can be modeled as a three-state monocyclic system. A single enzyme molecule can convert stochastically among three conformal states: the free enzyme E, the enzyme-substrate complex ES, and the enzyme-product complex EP. The product rate of the enzyme reaction is exactly the empirical net LE current of the monocyclic system. Using the exact joint distribution for LE currents, we find that rapid equilibrium between enzyme states E and ES and rapid equilibrium between enzyme states ES and EP both result in larger fluctuations in the product rate. Moreover, compared with the analytical solution, we show that the large deviations for LE currents provide a much more efficient computational method of the joint distribution, and thus enable a fast exploration of large swaths of parameter space.

Finally, we examine various types of fluctuation theorems satisfied by empirical LE and ST currents and clarify their ranges of applicability. We first show that the empirical absolute and net LE currents satisfy all types of fluctuation theorems and symmetric relations. In particular, we introduce the concept of similar cycles and obtain the strong form of the transient fluctuation theorem: (i) the joint distribution of empirical LE currents satisfies a symmetric relation when the currents of any pair of similar cycles are exchanged and (ii) the joint distribution of empirical net LE currents satisfies a symmetry relation when the net current of any cycle is replaced by its opposite number. Since empirical ST currents can be represented by the weight sums of empirical LE currents under the periodic boundary condition, we further show that empirical ST currents do not satisfy any form of fluctuation theorems, while empirical net ST currents only satisfy the weak form of the transient fluctuation theorem under the periodic boundary condition: the joint distribution of empirical net ST currents satisfies a symmetry relation when the net currents of all cycles in the fundamental set are replaced by their opposite numbers. As a corollary of the integral fluctuation theorem, we show that LE currents satisfy a refined version of the second law of thermodynamics: the entropy production along any finite number of cycles has a nonnegative mean, while it is broken for ST currents.

In the present paper, some results are only obtained for a monocyclic Markov chain. We anticipate that these results can be generalized to more general Markovian systems and even to semi-Markovian or non-Markovian systems. In addition, here we only make a comparison between LE and ST currents. The relationship between these two types of cycle currents and those defined in the sequence matching manner [18–20] is not clear. These are under current investigation.

ACKNOWLEDGMENTS

We are grateful to Profs. H. Qian and M. Qian for stimulating discussions. C.J. acknowledges support from National Natural Science Foundation of China with Grants No. U2230402, No. 12131005, and No. 12271020.

APPENDIX A: LARGE DEVIATIONS FOR LE CURRENTS OF A MONOCYCLIC MARKOV CHAIN

Here we give rigorous proof of the large deviation principle for LE currents. Recall that the empirical LE currents $(J_n^c)_{c\in\mathcal{C}}$ are said to satisfy a large deviation principle with a good rate function $I_J:\mathcal{V}\to[0,\infty]$ if [38]

(i) For each $\alpha\geqslant 0$, the level set $\{x\in\mathcal{V}: I_J(x)\leqslant\alpha\}$ is compact.

(ii) For each open set $U\subset\mathcal{V}$,

$$\varliminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\left(J_n^c\right)_{c\in\mathcal{C}}\in U\right)\geqslant-\inf_{x\in U}I_J(x). \tag{A1}$$

(iii) For each closed set $F\subset\mathcal{V}$,

$$\varlimsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\left(J_n^c\right)_{c\in\mathcal{C}}\in F\right)\leqslant-\inf_{x\in F}I_J(x). \tag{A2}$$

To prove the large deviation principle, we proceed in two steps. First, we examine the properties of the rate function and then we give the rigorous proof of upper bound Eq. (A1) and lower bound Eq. (A2).

Proposition 1. The rate function I_J is finite, continuous, and convex.

Proof. Recall that the rate function $I_J:\mathcal{V}\to[0,\infty]$ for empirical LE currents in the main text is given by

$$\begin{aligned}I_J(\nu)&=[h(\nu^{12})+h(\nu^{N1})+h(\nu^+)+h(\nu^-)-h(\nu^{12}+\nu^{N1}+\nu^++\nu^-)]\\&\quad+\inf_{X\in V(\nu)}F_\nu(X)+\sum_{i\in S}[h(\nu_i-\nu^i)+h(\nu^i)-h(\nu_i)]-\sum_{c\in\mathcal{C}}\nu^c\log\gamma^c\\&:=I_1(\nu)+I_2(\nu)+I_3(\nu)-I_4(\nu),\end{aligned} \tag{A3}$$

where $h(x)=x\log x$ is continuous on $[0,\infty)$ and

$$F_\nu(X)=\sum_{i=2}^{N-1}[h(x^i)+h(x^{i-1}+\nu^+)-h(x^i+x^{i-1}+\nu^+)]+\sum_{i=2}^{N-1}[h(y^i)+h(y^{i+1}+\nu^-)-h(y^i+y^{i+1}+\nu^-)]. \tag{A4}$$

It is easy to see that I_J is finite. We next prove that I_J is continuous. Obviously, I_1, I_3, and I_4 are continuous functions with respect to ν. Specially, I_4 is a linear function. We still need to prove I_2 is continuous. Let $Y(\nu)\in V(\nu)$ be a solution of Eq. (B2) in Appendix B. Since Eq. (B2) is a set of polynomial equations, it is easy to see that $Y(\nu)$ is a continuous function with respect to ν. This fact, together with the continuity of F_ν, guarantees that $\inf_{X\in V(\nu)}F_\nu(X)=F_\nu(Y(\nu))$ is a continuous function of ν.

Finally, we prove that I_J is convex. Note that

$$I_1(\nu)=\nu^{12}\log\left(\frac{\nu^{12}}{\hat{\nu}}\right)+\nu^{N1}\log\left(\frac{\nu^{N1}}{\hat{\nu}}\right)+\nu^+\log\left(\frac{\nu^+}{\hat{\nu}}\right)+\nu^-\log\left(\frac{\nu^-}{\hat{\nu}}\right), \tag{A5}$$

where $\hat{\nu} = \nu^{12} + \nu^{N1} + \nu^+ + \nu^-$. Recall the following log sum inequality: For any $a_1, a_2, b_1, b_2 \geqslant 0$, we have

$$(a_1 + a_2)\log\frac{a_1 + a_2}{b_1 + b_2} \leqslant a_1 \log\frac{a_1}{b_1} + a_2 \log\frac{a_2}{b_2}. \tag{A6}$$

For any $\alpha, \beta \geqslant 0$ satisfying $\alpha + \beta = 1$ and $\nu, \mu \in \mathcal{V}$, it follows from the log sum inequality Eq. (A6) that

$$(\alpha\nu^{12} + \beta\mu^{12})\log\left(\frac{\alpha\nu^{12} + \beta\mu^{12}}{\alpha\hat{\nu} + \beta\hat{\mu}}\right) \leqslant \alpha\nu^{12}\log\left(\frac{\nu^{12}}{\hat{\nu}}\right) + \beta\mu^{12}\log\left(\frac{\mu^{12}}{\hat{\mu}}\right),$$

where $\hat{\mu} = \mu^{12} + \mu^{N1} + \mu^+ + \mu^-$. This implies that the first term on the right-hand side of Eq. (A5) is a convex function of ν. Similarly, it is easy to prove that the other three terms on the right-hand side of Eq. (A5) are also convex functions of ν. This shows that I_1 is convex. Similarly, by the log sum inequality, we can also prove that I_3 is convex. Since I_4 is a linear function of ν, it is also convex. Finally, we will prove that $\inf_{X\in V(\nu)} F_\nu(X)$ is convex with respect to ν. To this end, we rewrite Eq. (A4) as

$$\begin{aligned}
F_\nu(X) &= \sum_{i=2}^{N-1}\left[(x^{i-1} + \nu^+)\log\left(\frac{x^{i-1} + \nu^+}{x^{i-1} + x^i + \nu^+}\right) + x^i \log\left(\frac{x^i}{x^{i-1} + x^i + \nu^+}\right)\right] \\
&\quad + \sum_{i=2}^{N-1}\left[y^i \log\left(\frac{y^i}{y^i + y^{i+1} + \nu^-}\right) + (y^{i+1} + \nu^-)\log\left(\frac{y^{i+1} + \nu^-}{y^i + y^{i+1} + \nu^-}\right)\right] \\
&:= \sum_{i=2}^{N-1}\left[A_1^i(\nu, X) + A_2^i(\nu, X) + A_3^i(\nu, X) + A_4^i(\nu, X)\right].
\end{aligned} \tag{A7}$$

Note that $\alpha X + \beta Z \in V(\alpha\nu + \beta\mu)$ for any $X = (x^i, y^i) \in V(\nu)$ and $Z = (z^i, w^i) \in V(\mu)$. Then by the log sum inequality, we have

$$\begin{aligned}
&A_1^i(\alpha\nu + \beta\mu, \alpha X + \beta Y) \\
&= (\alpha x^{i-1} + \beta z^{i-1} + \alpha\nu^+ + \beta\mu^+)\log\left(\frac{\alpha x^{i-1} + \beta z^{i-1} + \alpha\nu^+ + \beta\mu^+}{\alpha x^{i-1} + \alpha x^i + \beta z^{i-1} + \beta z^i + \alpha\nu^+ + \beta\mu^+}\right) \\
&= (\alpha(x^{i-1} + \nu^+) + \beta(z^{i-1} + \mu^+))\log\left(\frac{\alpha(x^{i-1} + \nu^+) + \beta(z^{i-1} + \mu^+)}{\alpha(x^{i-1} + x^i + \nu^+) + \beta(z^{i-1} + z^i + \mu^+)}\right) \\
&\leqslant \alpha(x^{i-1} + \nu^+)\log\left(\frac{x^{i-1} + \nu^+}{x^{i-1} + x^i + \nu^+}\right) + \beta(z^{i-1} + \mu^+)\log\left(\frac{z^{i-1} + \mu^+}{z^{i-1} + z^i + \mu^+}\right) \\
&= \alpha A_1^i(\nu, X) + \beta A_1^i(\mu, Y).
\end{aligned}$$

Similarly, we have $A_j^i(\alpha\nu + \beta\mu, \alpha X + \beta Y) \leqslant \alpha A_j^i(\nu, X) + \beta A_j^i(\mu, Y)$ for $j = 2, 3, 4$. This shows that

$$F_{\alpha\nu+\beta\mu}(\alpha X + \beta Y) \leqslant \alpha F_\nu(X) + \beta F_\mu(Y).$$

Optimizing over X and Y, we obtain

$$\inf_{Z\in V(\alpha\nu+\beta\mu)} F_{\alpha\nu+\beta\mu}(Z) \leqslant \alpha \inf_{X\in V(\nu)} F_\nu(X) + \beta \inf_{Y\in V(\mu)} F_\mu(Y).$$

This completes the proof of this proposition. ■

We next give the rigorous proof of the upper and lower bounds for empirical LE currents $(J_n^c)_{c\in\mathcal{C}}$.

Proposition 2. $(J_n^c)_{c\in\mathcal{C}}$ satisfy a large deviation principle with a good rate function I_J. Moreover, the upper bound of the large deviation principle can be improved as

$$\overline{\lim_{n\to+\infty}}\ \frac{1}{n}\log\mathbb{P}\left((J_n^c)_{c\in\mathcal{C}} \in \Gamma\right) \leqslant -\inf_{\nu\in\Gamma} I_J(\nu) \tag{A8}$$

for any set $\Gamma \subset \mathcal{V}$.

Proof. Since $\mathcal{V}$ is compact, the level set $\{x \in \mathcal{V} : I_J(x) \leqslant \alpha\}$ is also compact for any $\alpha \geqslant 0$. Set

$$K_n := \left\{(k^c)_{c\in\mathcal{C}} \in \mathbb{N}^{2N+2} : \sum_{c\in\mathcal{C}} k^c|c| = n\right\}. \tag{A9}$$

Without loss of generality, we assume that the system starts from state 1. For any $k = (k^c)_{c\in\mathcal{C}} \in K_n$, it follows from Eq. (7) in the main text that

$$\mathbb{P}_1\left(J_n^c = \frac{k^c}{n},\ \forall c \in \mathcal{C}\right) = |G_n(k)| \prod_{c\in\mathcal{C}} (\gamma^c)^{k^c}. \tag{A10}$$

Let $\mu_n(k) = k/n \in \mathcal{V}$. For any $\Gamma \subset \mathcal{V}$, let us put

$$Q_n(\Gamma) = \max_{k\in K_n:\mu_n(k)\in\Gamma} \mathbb{P}_1\left(J_n^c = \frac{k^c}{n},\ \forall c \in \mathcal{C}\right).$$

Then, clearly, we have

$$Q_n(\Gamma) \leqslant \mathbb{P}_1(J_n \in \Gamma) \leqslant |K_n| Q_n(\Gamma). \tag{A11}$$

It is easy to see that $|K_n| \leqslant (2N+2)(n+1)^{2N+3}$. Combining Eqs. (A10) and (A11), we obtain

$$\begin{aligned}\frac{1}{n}\log \mathbb{P}_1(J_n \in \Gamma) &= O\left(\frac{\log n}{n}\right) + \frac{1}{n}\log Q_n(\Gamma)\\ &= O\left(\frac{\log n}{n}\right) + \max_{k\in K_n:\mu_n(k)\in\Gamma}\left[\frac{1}{n}\log|G_n(k)| + \sum_{c\in\mathcal{C}}\frac{k_c}{n}\log\gamma^c\right]\\ &= O\left(\frac{\log n}{n}\right) - \min_{k\in K_n:\mu_n(k)\in\Gamma} I_J(\mu_n(k)).\end{aligned} \tag{A12}$$

To complete the proof, it is easy to see that $\cup_{n\in\mathbb{N}}\{\mu_n(k) : k \in K_n\}$ is dense in $\mathcal{V}$. This fact, together with the continuity of I_J, guarantees that for each $\nu \in \mathcal{V}$, there exists a sequence $(k_n)_{n\in\mathbb{N}}$ with $k_n \in K_n$ for all n, such that

$$\lim_{n\to\infty} \|\mu_n(k_n) - \nu\| = 0, \qquad \lim_{n\to\infty} I_J(\mu_n(k_n)) = I_J(\nu).$$

Then, for any open set $U \subset \mathcal{V}$, we have

$$\overline{\lim_{n\to\infty}} \min_{k\in K_n:\mu_n(k)\in U} I_J(\mu_n(k)) \leqslant I_J(\nu), \quad \forall \nu \in U.$$

Optimizing over $\nu \in U$, we obtain

$$\overline{\lim_{n\to\infty}} \min_{k\in K_n:\mu_n(k)\in U} I_J(\mu_n(k)) \leqslant \inf_{\nu\in U} I_J(\nu). \tag{A13}$$

Combining (A12) and (A13), we obtain the lower bound (A1) of the large deviation principle. Similarly, for any $\Gamma \subset \mathcal{V}$, we can prove that

$$\underline{\lim_{n\to\infty}} \min_{k\in K_n:\mu_n(k)\in\Gamma} I_J(\mu_n(k)) \geqslant \inf_{\nu\in\Gamma} I_J(\nu). \tag{A14}$$

Combining Eqs. (A12) and (A14), we obtain the upper bound Eq. (A2) of the large deviation principle. This completes the proof. ■

APPENDIX B: EXPLICIT EXPRESSION OF THE RATE FUNCTION I_J FOR MONOCYCLIC MARKOV CHAINS

Recall that for a monocyclic system, the rate function I_J for empirical LE currents is given by Eq. (15). Here the term $\inf_{X\in V(\nu)} F_\nu(X)$ is not in closed form. Next we will use the Lagrange multiplier method to find the explicit expression of $\inf_{X\in V(\nu)} F_\nu(X)$. For any $\nu \in \mathcal{V}$, we defined the Lagrangian function

$$\mathcal{A}_\nu(X,\lambda) = F_\nu(X) + \sum_{i=2}^{N-1}\lambda_i(x^i + y^i - \nu^{i,i+1}),$$

where $X = (x^i, y^i)_{2\leqslant i\leqslant N-1} \in V(\nu)$ and $\lambda = (\lambda_i)_{2\leqslant i\leqslant N-1} \in \mathbb{R}^{N-2}$. Taking the derivative of $\mathcal{A}_\nu(X,\lambda)$ with respect to the variables x^i, y^i, and λ_i, we obtain the following equations:

$$\begin{aligned}&\log(x^i) - \log(x^{i-1} + x^i + \nu^+) + \log(x^i + \nu^+) - \log(x^i + x^{i+1} + \nu^+) + \lambda_i = 0,\\ &\log(y^i + \nu^-) - \log(y^{i-1} + y^i + \nu^-) + \log(y^i) - \log(y^i + y^{i+1} + \nu^-) + \lambda_i = 0, \quad x^i + y^i = \nu^{i,i+1}, \quad 2 \leqslant i \leqslant N-1.\end{aligned} \tag{B1}$$

It is easy to check that Eqs. (B1) can be rewritten as

$$\frac{x^i}{x^{i-1} + x^i + \nu^+}\,\frac{x^i + \nu^+}{x^i + x^{i+1} + \nu^+} = \frac{y^i + \nu^-}{y^{i-1} + y^i + \nu^-}\,\frac{y^i}{y^i + y^{i+1} + \nu^-} = e^{-\lambda_i}, \quad x^i + y^i = \nu^{i,i+1}, \quad 2 \leqslant i \leqslant N-1. \tag{B2}$$

where $x^1 = \nu^{12}$, $x^N = 0$, $y^1 = 0$, and $y^N = \nu^{N1}$. Next we will prove that the set of algebraic equations Eqs. (B2) has at least one solution $X = (x^i, y^i) \in V(\nu)$ and this solution minimizes the function F_ν over $V(\nu)$.

Proposition 3. There exists at least one solution $X = (x^i, y^i) \in V(\nu)$ of Eqs. (B2).

Proof. If there exists $\nu^{k,k+1} = 0$ for some $2 \leqslant k \leqslant N-1$, then we have $x^k = y^k = 0$. Then Eqs. (B2) can be divided into two equations in terms of the indices $2 \leqslant i \leqslant k-1$ and $k+1 \leqslant i \leqslant N-1$. Thus, we only need to proof the lemma when $\nu^{k,k+1} > 0$ for any $2 \leqslant k \leqslant N-1$. Next we will consider three different cases.

Case 1: $\nu^{12} = \nu^+ = \nu^{N1} = \nu^- = 0$. It is easy to see that for each $\alpha \in (0, 1)$,

$$x^i = \alpha\nu^{i,i+1}, \quad y^i = (1-\alpha)\nu^{i,i+1}, \quad 2 \leqslant i \leqslant N-1,$$

is a solution of Eqs. (B2).

Case 2: $\nu^{12} = \nu^+ = 0$, $\nu^{N1} + \nu^- > 0$ or $\nu^{N1} = \nu^- = 0$, $\nu^{12} + \nu^+ > 0$. When $\nu^{12} = \nu^+ = 0$, $\nu^{N1} + \nu^- > 0$, it is easy to see that $x^i = 0$, $y^i = \nu^{i,i+1}$ is a solution of Eqs. (B2). Similarly, when $\nu^{N1} = \nu^- = 0$, $\nu^{12} + \nu^+ > 0$, it is easy to see that $x^i = \nu^{i,i+1}$, $y^i = 0$ is a solution of Eqs. (B2).

Case 3: $\nu^{12} + \nu^+ > 0$, $\nu^{N1} + \nu^- > 0$. We claim that for any $2 \leqslant k \leqslant N-2$, whenever $x^{k+1} \geqslant 0$, $y^{k+1} > 0$, and $x^{k+1} + y^{k+1} = \nu^{k+1,k+2}$, there exists a solution of the following set of equations:

$$\frac{x^i}{y^i} = \frac{x^i + x^{i+1} + \nu^+}{x^i + \nu^+}\,\frac{y^i + \nu^-}{y^{i-1} + y^i + \nu^-}\,\frac{x^{i-1} + x^i + \nu^+}{y^i + y^{i+1} + \nu^-}, \quad x^i + y^i = \nu^{i,i+1}, \quad 2 \leqslant i \leqslant k, \tag{B3}$$

and the solution satisfies $x^i, y^i > 0$ for all $2 \leqslant i \leqslant k$. We next prove it by induction. When $k = 2$, Eqs. (B3) can be simplified as

$$\frac{x^2}{\nu^{23} - x^2} = \frac{x^2 + x^3 + \nu^+}{x^2 + \nu^+}\,\frac{\nu^{12} + x^2 + \nu^+}{\nu^{23} - x^2 + y^3 + \nu^-}.$$

It is easy to see that

$$\lim_{x^2\downarrow 0} \frac{x^2}{\nu^{23} - x^2} = 0, \quad \lim_{x^2\downarrow 0} \frac{x^2 + x^3 + \nu^+}{x^2 + \nu^+}\,\frac{\nu^{12} + x^2 + \nu^+}{\nu^{23} - x^2 + y^3 + \nu^-} \geqslant \frac{\nu^{12} + \nu^+}{\nu^{23} + y^3 + \nu^-} > 0.$$

On the other hand, we have

$$\lim_{x^2\uparrow \nu^{23}} \frac{x^2}{\nu^{23} - x^2} = \infty, \quad \lim_{x^2\uparrow \nu^{23}} \frac{x^2 + x^3 + \nu^+}{x^2 + \nu^+}\,\frac{\nu^{12} + x^2 + \nu^+}{\nu^{23} - x^2 + y^3 + \nu^-} = \frac{\nu^{23} + x^3 + \nu^+}{\nu^{23} + \nu^+}\,\frac{\nu^{12} + \nu^{23} + \nu^+}{y^3 + \nu^-} < \infty.$$

By the intermediate value theorem, we can find a solution of Eqs. (B3) satisfying $x^2, y^2 > 0$ and $x^2 + y^2 = \nu^{23}$. Suppose that the claim holds for $k = n-1$. Then we consider the equation

$$\frac{x^n}{\nu^{n,n+1} - x^n} = \frac{x^n + x^{n+1} + \nu^+}{x^n + \nu^+}\,\frac{\nu^{n,n+1} - x^n + \nu^-}{y^{n-1} + \nu^{n,n+1} - x^n + \nu^-}\,\frac{x^{n-1} + x^n + \nu^+}{\nu^{n,n+1} - x^n + y^{n+1} + \nu^-},$$

where $x^{n-1}, y^{n-1} > 0$ is the solution of Eqs. (B3) for $k = n-1$, which depends on $x^n \geqslant 0$. It is easy to see that

$$\lim_{x^n\downarrow 0} \frac{x^n}{\nu^{n,n+1} - x^n} = 0, \tag{B4}$$

and

$$\begin{aligned}
&\lim_{x^n\downarrow 0} \frac{x^n + x^{n+1} + \nu^+}{x^n + \nu^+}\,\frac{\nu^{n,n+1} - x^n + \nu^-}{y^{n-1} + \nu^{n,n+1} - x^n + \nu^-}\,\frac{x^{n-1} + x^n + \nu^+}{\nu^{n,n+1} - x^n + y^{n+1} + \nu^-}\\
&\geqslant \frac{\nu^{n,n+1} + \nu^-}{\lim_{x^n\downarrow 0} y^{n-1} + \nu^{n,n+1} + \nu^-}\,\frac{\lim_{x^n\downarrow 0} x^{n-1} + \nu^+}{\nu^{n,n+1} + y^{n+1} + \nu^-} > 0,
\end{aligned} \tag{B5}$$

where we have used the fact that $\lim_{x^n\downarrow 0} x^{n-1} > 0$ since x^{n-1} is a continuous function of x^n. On the other hand, we have

$$\lim_{x^n\uparrow \nu^{n,n+1}} \frac{x^n}{\nu^{n,n+1} - x^n} = \infty \tag{B6}$$

and

$$\begin{aligned}
&\lim_{x^n\uparrow \nu^{n,n+1}} \frac{x^n + x^{n+1} + \nu^+}{x^n + \nu^+}\,\frac{\nu^{n,n+1} - x^n + \nu^-}{y^{n-1} + \nu^{n,n+1} - x^n + \nu^-}\,\frac{x^{n-1} + x^n + \nu^+}{\nu^{n,n+1} - x^n + y^{n+1} + \nu^-}\\
&\leqslant \frac{\nu^{n,n+1} + x^{n+1} + \nu^+}{\nu^{n,n+1} + \nu^+}\,\frac{\lim_{x^n\uparrow \nu^{n,n+1}} x^{n-1} + \nu^{n,n+1} + \nu^+}{y^{n+1} + \nu^-} < \infty.
\end{aligned} \tag{B7}$$

By the intermediate value theorem, we can find a solution of Eqs. (B3) satisfying $x^i, y^i > 0$ and $x^i + y^i = \nu^{i,i+1}$ for any $2 \leqslant i \leqslant n$. By induction, the claim holds for all $k \geqslant 2$.

Note that when $k = N-1$, Eqs. (B3) are equivalent to Eqs. (B2). In this case, we have $y^{k+1} = \nu^{N1}$. Since $\nu^{N1} + \nu^- > 0$, it is easy to see that Eqs. (B4)–(B7) still hold. Similarly to the above proof, by the intermediate value theorem, we can find a solution of Eqs. (B3) satisfying $x^i, y^i > 0$ and $x^i + y^i = \nu^{i,i+1}$ for any $2 \leqslant i \leqslant N-1$. This completes the proof of this lemma. ■

Proposition 4. Let $X = (x^i, y^i) \in V(\nu)$ be any solution of Eqs. (B2). Then X minimizes the function F_ν over $V(\nu)$.

Proof. By the log sum inequality Eq. (A6), we have

$$a_1 \log \frac{a_1}{a_1 + a_2} + a_2 \log \frac{a_2}{a_1 + a_2} \geqslant a_1 \log \frac{b_1}{b_1 + b_2} + a_2 \log \frac{b_2}{b_1 + b_2}. \tag{B8}$$

For any $Z = (z^i, w^i) \in V(\nu)$, it follows from Eq. (A7) that

$$F_\nu(Z) = \sum_{i=2}^{N-1} \left[(z^{i-1} + \nu^+) \log \frac{z^{i-1} + \nu^+}{z^{i-1} + z^i + \nu^+} + z^i \log \frac{z^i}{z^{i-1} + z^i + \nu^+} \right] + \sum_{i=2}^{N-1} \left[w^i \log \frac{w^i}{w^i + w^{i+1} + \nu^-} + (w^{i+1} + \nu^-) \log \frac{w^{i+1} + \nu^-}{w^i + w^{i+1} + \nu^-} \right],$$

where $z^1 = \nu^{12}$, $w^N = \nu^{N1}$. Combining Eqs. (B2) and (B8), we have

$$\begin{aligned}
F_\nu(Z) &\geqslant \sum_{i=2}^{N-1} \left[(z^{i-1} + \nu^+) \log \frac{x^{i-1} + \nu^+}{x^{i-1} + x^i + \nu^+} + z^i \log \frac{x^i}{x^{i-1} + x^i + \nu^+} \right] \\
&\quad + \sum_{i=2}^{N-1} \left[w^i \log \frac{y^i}{y^i + y^{i+1} + \nu^-} + (w^{i+1} + \nu^-) \log \frac{y^{i+1} + \nu^-}{y^i + y^{i+1} + \nu^-} \right] \\
&= \sum_{i=2}^{N-1} \left[\nu^+ \log \frac{x^{i-1} + \nu^+}{x^{i-1} + x^i + \nu^+} + z^i \log \frac{x^i + \nu^+}{x^{i-1} + x^i + \nu^+} \frac{x^i}{x^{i-1} + x^i + \nu^+} \right] \\
&\quad + \sum_{i=2}^{N-1} \left[\nu^- \log \frac{y^{i+1} + \nu^-}{y^i + y^{i+1} + \nu^-} + w^i \log \frac{y^i + \nu^-}{y^{i-1} + y^i + \nu^-} \frac{y^i}{y^i + y^{i+1} + \nu^-} \right] \\
&\quad + \nu^{12} \log \frac{\nu^{12} + \nu^+}{\nu^{12} + x^2 + \nu^+} + \nu^{N1} \log \frac{\nu^{N1} + \nu^-}{\nu^{N1} + y^{N-1} + \nu^-} \\
&= \sum_{i=2}^{N-1} \left[-\lambda_i \nu^{i,i+1} + \nu^+ \log \frac{x^{i-1} + \nu^+}{x^{i-1} + x^i + \nu^+} + \nu^- \log \frac{y^{i+1} + \nu^-}{y^i + y^{i+1} + \nu^-} \right] \\
&\quad + \nu^{12} \log \frac{\nu^{12} + \nu^+}{\nu^{12} + x^2 + \nu^+} + \nu^{N1} \log \frac{\nu^{N1} + \nu^-}{\nu^{N1} + y^{N-1} + \nu^-} \\
&= \sum_{i=2}^{N-1} \left[(x^{i-1} + \nu^+) \log \frac{x^{i-1} + \nu^+}{x^{i-1} + x^i + \nu^+} + x^i \log \frac{x^i}{x^{i-1} + x^i + \nu^+} \right] \\
&\quad + \sum_{i=2}^{N-1} \left[y^i \log \frac{y^i}{y^i + y^{i+1} + \nu^-} + (y^{i+1} + \nu^-) \log \frac{y^{i+1} + \nu^-}{y^i + y^{i+1} + \nu^-} \right] \\
&= F_\nu(X),
\end{aligned}$$

where λ_i is defined in Eqs. (B2). This gives the desired result. ∎

APPENDIX C: INDEPENDENCE OF THE RATE FUNCTION WITH RESPECT TO THE INITIAL DISTRIBUTION

Here we will prove that for any monocyclic system, the rate function I_J is independent of the choice of the initial distribution. Before giving the proof, we introduce some notations.

For any $k = (k^c)_{c \in \mathcal{C}} \in \mathbb{N}^{2N+2}$, set $|k| = \sum_{c \in \mathcal{C}} k^c |c|$. For any state $i \in S$, let $G^i(k)$ be the set of all possible trajectories (with periodic boundary conditions) up to time n with initial state i so each cycle c is formed k^c times, i.e.,

$$G^i(k) = \left\{ (\xi_0, \xi_1, \cdots, \xi_n) \in S^{n+1} : N_n^c = k^c \text{ for any } \forall c \in \mathcal{C},\ \xi_0 = i,\ n = |k| \right\}.$$

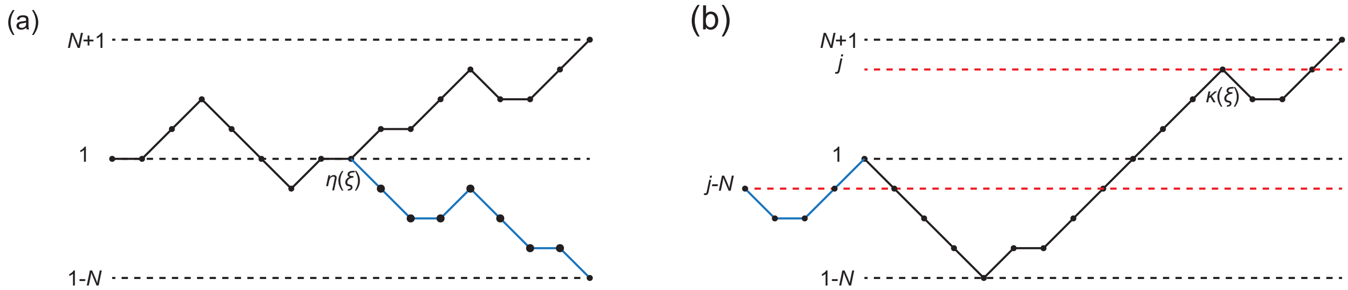

FIG. 5. The broken line graph for the trajectory ξ. Here we consider states i and $N+i$ as different states. (a) The first type mapping of trajectory: $\eta(\xi)$ is the last time that trajectory ξ reach state 1. (b) The second type mapping of trajectory: $\kappa(\xi)$ denotes the first time that trajectory ξ reaches state j.

For any state $i \in S$ and any cycle $c \ni i$, let $G^{i,c}(k)$ be the subset of $G^i(k)$ so the last cycle formed by the trajectory is cycle c. For convenience, in the following, if $c = C^+$, then $G^{i,c}(k)$ is abbreviated as $G^{i,+}(k)$; if $c = C^-$, then $G^{i,c}(k)$ is abbreviated as $G^{i,-}(k)$.

We next give some nontrivial identities about $|G^i(k)|$ and $|G^{i,c}(k)|$.

Lemma 1. Set

$$B^i(\tilde{k}) = |G^{i,+}(1,0,\tilde{k})|, \quad C^i(\tilde{k}) = |G^{i,-}(0,1,\tilde{k})|,$$

where $\tilde{k} = (k^1,\cdots,k^N,k^{12},\cdots,k^{N1})$. Then, for any $i, j \in S$ and $\tilde{k} \in \mathbb{N}^{2N}$, we have

$$B^i(\tilde{k}) = B^j(\tilde{k}) = C^i(\tilde{k}) = C^j(\tilde{k}).$$

Proof. Without loss of generality, we assume $i = 1$. We first prove that $B^1(\tilde{k}) = C^1(\tilde{k})$. Note that any monocyclic Markov chain naturally corresponds to a nearest-neighbor random walk (with the transition from a state to itself being allowed). Hence, each trajectory of the monocyclic system can be represented as a broken line graph illustrated in Fig. 5. The horizontal movement of the broken line from state i to itself corresponds to the formation of cycle (i). Each decrease of the broken line from state $i+1$ to state i above state 1 corresponds to the formation of cycle $(i, i+1)$. Each increase of the broken line from state $-j$ to state $1-j$ below state 1 corresponds to the formation of cycle $(-j, 1-j)$, where all states are understood to be modulo N.

Let $n = N + |\tilde{k}|$. For any trajectory $\xi = (\xi_m)_{0\leqslant m\leqslant n} \in G^{1,+}(1,0,\tilde{k})$, let

$$\eta(\xi) = \max\{m : 0 \leqslant m \leqslant n-1, \xi_m = 1\}$$

be the last exit time of the initial state 1 before time n. Since the last cycle formed by ξ is C^+, it is clear that ξ must reach the set $\{N+1, 1-N\}$ for the first time at time n with $\xi_n = N+1$. Then we can construct another trajectory $\tilde{\xi} = (\tilde{\xi}_m)_{0\leqslant m\leqslant n}$ corresponding to ξ as [see Fig. 5(a) for an illustration]

$$\tilde{\xi}_m := \begin{cases} \xi_m, & \text{if } 0 \leqslant m \leqslant \eta(\xi) \\ \xi_{n+\eta(\xi)-m} - N, & \text{if } \eta(\xi) < m \leqslant n, \end{cases}$$

where the original trajectory between time $\eta(\xi)$ and time n are reversed and then translated to the initial state 1. Clearly, each cycle formed by $(\xi_m)_{0\leqslant m\leqslant \eta(\xi)}$ is also formed by $(\tilde{\xi}_m)_{0\leqslant m\leqslant \eta(\xi)}$. In addition, the horizontal movement of the broken line from state i to itself for $(\xi_m)_{\eta(\xi)\leqslant m\leqslant n}$ corresponds to the horizontal movement of the broken line from state $i-N$ to itself for $(\tilde{\xi}_m)_{\eta(\xi)\leqslant m\leqslant n}$. Similarly, each decrease of the broken line from state $i+1$ to state i for $(\xi_m)_{\eta(\xi)\leqslant m\leqslant n}$ corresponds to an increase of the broken line from state $i-N$ to state $i+1-N$ for $(\tilde{\xi}_m)_{\eta(\xi)\leqslant m\leqslant n}$. Then all one-state and two-state cycles formed by ξ are also formed by $\tilde{\xi}$. Moreover, it is clear that $\tilde{\xi}$ reaches the set $\{N+1, 1-N\}$ for the first time at time n with $\tilde{\xi}_n = 1-N$. This implies that $\tilde{\xi} \in G^{1,-}(0,1,\tilde{k})$ since the last cycle formed by $\tilde{\xi}$ is C^-. Hence the map $\xi \longmapsto \tilde{\xi}$ gives a one-to-one correspondence between $G^{1,+}(1,0,\tilde{k})$ and $G^{1,-}(0,1,\tilde{k})$. This shows that $B^1(\tilde{k}) = C^1(\tilde{k})$.

We next prove that $B^1(\tilde{k}) = B^j(\tilde{k})$. For each $\tilde{k}$ satisfying $|\tilde{k}| = n - N$, let $\tilde{G}^l(\tilde{k})$ be the set of all possible trajectories up to time n with initial state l so each edge $\langle i,i\rangle$ is passed through k^i times, each edge $\langle i,i+1\rangle$ is passed through $k^{i,i+1}+1$ times, each edge $\langle i+1,i\rangle$ is passed through $k^{i,i+1}$ times, and the trajectory reaches state $N+i$ for the first time at time n, where all states are understood to be module N. We claim that $|\tilde{G}^1(\tilde{k})| = |\tilde{G}^j(\tilde{k})|$.

We now prove the claim. For any trajectory $\xi = (\xi_m)_{0\leqslant m\leqslant n} \in \tilde{G}^1(\tilde{k})$, let

$$\kappa^j(\xi) = \min\{m : 0 \leqslant m \leqslant n, \xi_m = j\}$$

be the hitting time of state j. Then we can construct another trajectory $\tilde{\xi} = (\tilde{\xi}_m)_{0\leqslant m\leqslant n}$ corresponding to ξ as [see Fig. 5(b) for an illustration]

$$\tilde{\xi}_m := \begin{cases} \xi_{\kappa^j(\xi)+m}, & \text{if } 0 \leqslant m \leqslant n-\kappa^j(\xi) \\ \xi_{m+\kappa^j(\xi)-n}, & \text{if } n-\kappa^j(\xi) < m \leqslant n, \end{cases}$$

where the original trajectory between time $\kappa^j(\xi)$ and time n are moved before the initial time and then translated to initial state 1. It is easy to see that $\tilde{\xi}$ has initial state j. In addition, edge $\langle i,i\rangle$ is passed through k^i times, edge $\langle i,i+1\rangle$ is passed through $k^{i,i+1}+1$ times and edge $\langle i+1,i\rangle$ are passed through $k^{i,i+1}$ times. Moreover, since ξ reaches state j is the first time at time $\kappa^j(\xi)$, it is clear that $\tilde{\xi}$ reaches state $j+N$ the first time at time n. This implies that $\tilde{\xi} \in \tilde{G}^j(\tilde{k})$. Hence the map $\xi \to \tilde{\xi}$ gives a one-to-one correspondence between $\tilde{G}^1(\tilde{k})$ and $\tilde{G}^j(\tilde{k})$. This shows that $|\tilde{G}^1(\tilde{k})| = |\tilde{G}^j(\tilde{k})|$.

In fact, for each trajectory $\xi \in \tilde{G}^1(\tilde{k})$, cycles (i), $(i,i+1)$, $(1,\cdots,N)$, and $(1,N,\cdots,2)$ are formed k^i, $k^{i,i+1}-r$, $r+1$, and r times, respectively. Here $r \leqslant \min_{i\in S} k^{i,i+1}$ depends on ξ. It is easy to see that ξ reaches state $N+1$ for the first time at time n. Then the last cycle formed by ξ is C^+ and thus

$$\tilde{G}^1(\tilde{k}) = \bigsqcup_{r=0}^{\min_{i\in S} k^{i,i+1}} G^{1,+}(r+1, r, k^1, \cdots, k^N, k^{12} - r, \cdots, k^{N1} - r).$$

Since $B^1(\tilde{k}) = |G^{1,+}(1, 0, \tilde{k})|$, we obtain

$$|\tilde{G}^1(\tilde{k})| = \sum_{r=1}^{\min_{i\in S} k^{i,i+1}} |G^{1,+}(r+1, r, k^1, \cdots, k^N, k^{12} - r, \cdots, k^{N1} - r)| + B^1(\tilde{k}). \tag{C1}$$

Similarly, we can also obtain

$$|\tilde{G}^j(\tilde{k})| = \sum_{r=1}^{\min_{i\in S} k^{i,i+1}} |G^{j,+}(r+1, r, k^1, \cdots, k^N, k^{12} - r, \cdots, k^{N1} - r)| + B^j(\tilde{k}). \tag{C2}$$

Finally, we prove $B^1(\tilde{k}) = B^j(\tilde{k})$ by induction. When $\min_{i\in S} k^{i,i+1} = 0$, it follows from Eqs. (C1), (C2), and the claim that

$$B^1(\tilde{k}) = |\tilde{G}^1(\tilde{k})| = |\tilde{G}^j(\tilde{k})| = B^j(\tilde{k}).$$

Suppose that the equality holds for $\min_{i\in S} k^{i,i+1} \leqslant t$. When $\min_{i\in S} k^{i,i+1} = t+1$, fix $1 \leqslant r \leqslant t+1$. Now we calculate $|G^{1,+}(r+1, r, k^1, \cdots, k^N, k^{12} - r, \cdots, k^{N1} - r)|$. Let $\xi \in G^{1,+}(r+1, r, k^1, \cdots, k^N, k^{12} - r, \cdots, k^{N1} - r)$, then trajectory ξ can be divided into $2r+1$ subtrajectories such that the last cycle formed by each subtrajectory is cycle $(1, \cdots, N)$ or $(1, N, \cdots, 2)$. Note that the number of permutations for inserting $r+1$ cycle $(1, \cdots, N)$ and r cycle $(1, N, \cdots, 2)$ into state 1 such that the last cycle is $(1, \cdots, N)$ is $\binom{2r}{r}$. Fix the partition of the rest cycles, i.e.,

$$\sum_{s=1}^{2r+1} \tilde{k}_s = (k^1, \cdots, k^N, k^{12} - r, \cdots, k^{N1} - r).$$

Since $B^i(\tilde{k}) = C^i(\tilde{k})$ for all i and $\tilde{k}$, the number of insertions is $\prod_{s=1}^{2r+1} B^j(\tilde{k}_s)$. Summing over all choices of $\tilde{k}_s$, the total number is given by

$$|G^{1,+}(r+1, r, k^1, \cdots, k^N, k^{12} - r, \cdots, k^{N1} - r)| = \binom{2r}{r} \sum_{\sum_{s=1}^{2r+1} \tilde{k}_s = (k^1, \cdots, k^N, k^{12}-r, \cdots, k^{N1}-r)} \prod_{s=1}^{2r+1} B^1(\tilde{k}_s).$$

Similarly, we can also obtain

$$|G^{j,+}(r+1, r, k^1, \cdots, k^N, k^{12} - r, \cdots, k^{N1} - r)| = \binom{2r}{r} \sum_{\sum_{s=1}^{2r+1} \tilde{k}_s = (k^1, \cdots, k^N, k^{12}-r, \cdots, k^{N1}-r)} \prod_{s=1}^{2r+1} B^j(\tilde{k}_s).$$

Note that $\min_{i\in S}(k^{i,i+1} - r) \leqslant \min_{i\in S} k^{i,i+1} - 1 \leqslant t$, we have $B^1(\tilde{k}_s) = B^j(\tilde{k}_s)$ by induction. Then

$$|G^{1,+}(r+1, r, k^1, \cdots, k^N, k^{12} - r, \cdots, k^{N1} - r)| = |G^{j,+}(r+1, r, k^1, \cdots, k^N, k^{12} - r, \cdots, k^{N1} - r)|.$$

It follows from Eqs. (C1), (C2), and the claim that $B^1(\tilde{k}) = B^j(\tilde{k})$. By induction, we complete the proof. ■

Proposition 5. For any $i, j \in S$, $k \in \mathbb{N}^{2N+2}$:

$$|G^{i,+}(k)| = |G^{j,+}(k)|, \quad |G^{i,-}(k)| = |G^{j,-}(k)|. \tag{C3}$$

Proof. Here we only prove the first identity of Eqs. (C3); the proof the second identity is similar. Now we calculate $|G^{i,+}(k)|$. Suppose that the system starts from state i. We first insert cycles C^+ and C^- into the trajectory. Note that the last cycle formed by any trajectory $\xi \in G_n^{i,+}(k)$ is C^+. Since cycle C^+ is formed k^+ times and cycle C^+ is formed k^+ times, the number of insertions is $\binom{k^+ + k^- - 1}{k^+ - 1}$.

We next insert the remaining one-state and two-state cycles into the trajectory. Since there are currently $k^+ + k^-$ N-state cycles in the trajectory, we divide the remaining cycles into $k^+ + k^-$ groups and then insert those cycles that belong to the sth group onto the sth N-state cycle. If we partition the remaining cycles into $k^+ + k^-$ groups as $\sum_{s=1}^{k^+ + k^-} \tilde{k}_s = \tilde{k}$, then it follows from Lemma 1 that the number of insertions is $\prod_{s=1}^{k^+ + k^-} B^i(\tilde{k}_s)$. Summing over all choices of $\tilde{k}_s$, the total number of insertions is given by

$$|G^{i,+}(k)| = \binom{k^+ + k^- - 1}{k^+ - 1} \sum_{\sum_{s=1}^{k^+ + k^-} \tilde{k}_s = \tilde{k}} \prod_{s=1}^{k^+ + k^-} B^i(\tilde{k}_s).$$

Similarly, we have

$$|G^{j,+}(k)| = \binom{k^+ + k^- - 1}{k^+ - 1} \sum_{\sum_{s=1}^{k^+ + k^-} \tilde{k}_s = \tilde{k}} \prod_{s=1}^{k^+ + k^-} B^j(\tilde{k}_s).$$

By Lemma 1, we complete the proof of this proposition. ■

The next lemma gives the relationship between $|G^i(k)|$ and $|G^{i,c}(k)|$.

Lemma 2. Let $k \in \mathbb{N}^{2N+2}$, $i \in S$, and $c \in \mathcal{C}$. Suppose that c passes through state i. Then

$$|G^{i,c}(k)| = \frac{k^c}{\sum_{c \ni i} k^c} |G^i(k)|.$$

Proof. Without loss of generality, we assume $i = 1$. Recall the three-step cycle insertion method of calculating $|G^1(k)|$ in the main text. Let us fix the last cycle as c in the first step and the number of permutations is given by

$$\frac{k^c}{k^1 + k^{12} + k^{N1} + k^+ + k^-} A_1.$$

Keep the second and third steps the same. In this case, the last cycle formed by the trajectory will be c, which means

$$|G^{1,c}(k)| = \frac{k^c}{\sum_{c \ni 1} k^c} |G^1(k)|.$$

This completes the proof of this lemma. ■

Proposition 6. Under the periodic boundary condition, the rate function I_J is independent of the choice of the initial distribution of ξ.

Proof. We only need to prove that the rate function when the systems starts from state 1 is exactly the same as the the rate function when the systems starts from state 2. For any state $i \in S$, recall that

$$\mathbb{P}_i\big(J_n^c = \nu^c, \forall\, c \in \mathcal{C}\big) = |G^i(k)| \prod_{c\in\mathcal{C}} (\gamma^c)^{k^c},$$

where $n = \sum_{c\in\mathcal{C}} |c|k^c$ and $\nu^c = k^c/n$. It follows from Eqs. (6) in the main text that we only need to prove that $\log|G^1(k)| = \log|G^2(k)| + O(\log n)$. Let $c = (1, \cdots, N)$. By Proposition 5 and Lemma 2, we have

$$\begin{aligned}\log|G^1(k)| &= \log|G^{1,+}(k)| + O(\log n)\\ &= \log|G^{2,+}(k)| + O(\log n)\\ &= \log|G^2(k)| + O(\log n).\end{aligned}$$

This completes the proof of this proposition. ■

APPENDIX D: SIMPLIFIED EXPRESSION OF RATE FUNCTION I_J IN TWO SPECIAL CASES

We have seen that the rate function I_J for empirical LE currents of a monocyclic system can be simplified to a large extent in two special cases: (i) the case where the system has only three states and (ii) the case where the transition from state 1 to state N is forbidden [see Fig. 1(d) for an illustration]. Next we will give the proof.

We first prove that for a three-state system, the rate function is given by Eq. (17). When $N = 3$, it is easy to see that the solution $X = (x^2, y^2)$ of Eqs. (16) is given by

$$x^2 = \frac{\nu^{23}(\nu^{12}+\nu^+)}{\nu^{12}+\nu^{13}+\nu^++\nu^-}, \quad y^2 = \frac{\nu^{23}(\nu^{13}+\nu^-)}{\nu^{12}+\nu^{13}+\nu^++\nu^-}.$$

Note that the solution $X = (x^2, y^2)$ minimizes the function F_ν. Then we have

$$\begin{aligned}I_2(\nu) &= F_\nu(X)\\ &= \nu^{23}\log\frac{\nu^{23}}{\tilde{\nu}} + (\nu^{12}+\nu^{13}+\nu^++\nu^-)\log\frac{\tilde{\nu}-\nu^{23}}{\tilde{\nu}}.\end{aligned}$$

Straightforward calculations show that

$$\begin{aligned}&I_1(\nu) + I_2(\nu) + I_3(\nu)\\ &\quad= \sum_{i\in S}\left[\nu^i\log\frac{\nu^i}{\nu_i} + (\nu_i - \nu^i)\log\frac{\nu_i-\nu^i}{\nu_i}\right]\\ &\qquad+ \sum_{c\in\mathcal{C},|c|\neq 1}\nu^c\log\frac{\nu^c}{\tilde{\nu}}. \end{aligned} \tag{D1}$$

Recall the following expression of the LE currents Ref. [17, Theorem.1.3.3],

$$\begin{aligned}J^+ &= \gamma^+\frac{1}{C}, \quad J^- = \gamma^-\frac{1}{C},\\ J^{i,i+1} &= \gamma^{i,i+1}\frac{1-p_{i-1,i-1}}{C}, \quad 1 \leqslant i \leqslant 3,\end{aligned} \tag{D2}$$

where $C = \sum_{i\in S}[(1-p_{i-1,i-1})(1-p_{i+1,i+1}) - p_{i-1,i+1}\, p_{i+1,i-1}]$. It then follows from Eq. (1) that

$$p_{ij} = \frac{\sum_{c\ni\langle i,j\rangle} J^c}{\sum_{c\ni i} J^c}. \tag{D3}$$

Combining Eqs. (D2) and (D3), we have

$$\begin{aligned}&\sum_{i\in S}\left[\nu^i\log\frac{J^i}{J_i} + (\nu_i-\nu^i)\log\frac{J_i-J^i}{J_i}\right] + \sum_{c\in\mathcal{C},|c|\neq1}\nu^c\log\left(\frac{J^c}{\tilde{J}}\right)\\ &= \sum_{i\in S}\left[\nu^i\log\frac{J^i}{J_i} + \nu^{i,i+1}\log\left(\left(1-\frac{J^i}{J_i}\right)\left(1-\frac{J^{i+1}}{J_{i+1}}\right)\frac{J^{i,i+1}}{\tilde{J}}\right)\right] + \nu^+\log\left(\frac{J^+}{\tilde{J}}\prod_{i\in\mathcal{C}}\left(1-\frac{J^i}{J_i}\right)\right) + \nu^-\log\left(\frac{J^-}{\tilde{J}}\prod_{i\in\mathcal{C}}\left(1-\frac{J^i}{J_i}\right)\right)\\ &= \sum_{c\in\mathcal{C}}\nu^c\log\gamma^c = -I_4(\nu).\end{aligned} \tag{D4}$$

Combining Eqs. (D1) and (D4) gives the desired result.

We next prove that for a monocyclic system, if the transition from state 1 to state N is forbidden [see Fig. 1(d) for an illustration], then the rate function is given by Eq. (18). Since $p_{N1} = 0$, the two cycles $(1, N)$ and $(1, N, \cdots, 2)$ cannot be formed. Hence we can take $\nu^{N1} = \nu^- = 0$ in Eqs. (16) and it is easy to see that $x^i = \nu^{i,i+1}$, $y^i = 0$ is a solution of Eqs. (16). Then we have

$$I_2(\nu) = F_\nu(x^i, y^i) = \sum_{i=2}^{N-1}\left[-\lambda_i\nu^{i,i+1} + \nu^+\log\frac{\nu^{i-1,i}+\nu^+}{\nu^{i-1,i}+\nu^{i,i+1}+\nu^+}\right] + \nu^{12}\log\frac{\nu^{12}+\nu^+}{\nu^{12}+\nu^{23}+\nu^+},$$

where

$$\lambda_i = -\log\left(\frac{\nu^{i,i+1}}{\nu^{i-1,i}+\nu^{i,i+1}+\nu^+}\,\frac{\nu^{i,i+1}+\nu^+}{\nu^{i,i+1}+\nu^{i+1,i+2}+\nu^+}\right).$$

By the definition of ν_i, we have

$$\nu_1 = \nu^1 + \nu^{12} + \nu^+, \quad \nu_i = \nu^i + \nu^{i-1,i} + \nu^{i,i+1} + \nu^+, \quad 2 \leqslant i \leqslant N-1, \quad \nu_N = \nu^N + \nu^{N-1,N} + \nu^+.$$

Straightforward calculations show that

$$
\begin{aligned}
I_1(\nu) &= \nu^{12}\log\frac{\nu^{12}}{\nu_1-\nu^{12}}+\nu^+\log\frac{\nu^+}{\nu_1-\nu^1},\\
I_2(\nu) &= \sum_{i=2}^N\left[\nu^{i,i+1}\log\frac{\nu^{i,i+1}}{\nu_i-\nu^i}+\nu^+\log\frac{\nu^{i-1,i}+\nu^+}{\nu_i-\nu^i}\right]+\sum_{i=1}^N\nu^{i,i+1}\log\frac{\nu^{i,i+1}+\nu^+}{\nu_{i+1}-\nu^{i+1}},\\
I_3(\nu) &= \sum_{i\in S}\left[\nu^i\log\frac{\nu^i}{\nu_i}+\nu^+\log\frac{\nu_i-\nu^i}{\nu_i}\right]+\sum_{i\in S}\nu^{i,i+1}\left(\log\frac{\nu_i-\nu^i}{\nu_i}+\log\frac{\nu_{i+1}-\nu^{i+1}}{\nu_{i+1}}\right).
\end{aligned}\tag{D5}
$$

It then follows from Eq. (D3) that

$$
\begin{aligned}
&\sum_{i\in S}\left[\nu^i\log\frac{J^i}{J_i}+\nu^{i,i+1}\log\frac{J^{i,i+1}}{J_i}+(\nu^{i-1,i}+\nu^+)\log\frac{J^{i-1,i}+J^+}{J_i}\right]\\
&\quad=\sum_{i\in S}\left[\nu^i\log\frac{J^i}{J_i}+\nu^{i,i+1}\log\frac{J^{i,i+1}(J^{i,i+1}+J^+)}{J_{i+1}J_i}\right]+\nu^+\log\frac{\prod_{i=1}^N(J^{i,i+1}+J^+)}{\prod_{i=1}^N J_i}\\
&\quad=\sum_{c\in\mathcal{C}}\nu^c\log\gamma^c=-I_4(\nu).
\end{aligned}\tag{D6}
$$

Combining Eqs. (D5) and (D6) gives the desired result.

APPENDIX E: PROOF OF THE TRANSIENT FLUCTUATION THEOREM FOR NET ST CURRENTS

Here we will prove Eq. (39) under the periodic boundary condition. It follows from Eq. (38) that for any cycle $c_l\in\mathcal{L}$ with three or more states, we have

$$\tilde{Q}_n^{c_l}=\sum_{i=1}^r\tilde{J}_n^{c_i}[1_{\{l\in c_i\}}-1_{\{l\in c_i-\}}],$$

where 1_A is the indicator function which takes the value of 1 when A holds and takes the value of 0 when A does not hold. Then we obtain

$$
\begin{aligned}
\mathbb{P}\left(\tilde{Q}_n^{c_{l_1}}=x_1,\cdots,\tilde{Q}_n^{c_{l_s}}=x_s\right) &= \mathbb{P}\left(\sum_{i=1}^r\tilde{J}_n^{c_i}\left[1_{\{l_1\in c_i\}}-1_{\{l_1\in c_i-\}}\right]=x_1,\cdots,\sum_{i=1}^r\tilde{J}_n^{c_i}\left[1_{\{l_s\in c_i\}}-1_{\{l_s\in c_i-\}}\right]=x_s\right)\\
&= \sum_{\sum_{i=1}^r y_i[1_{\{l_m\in c_i\}}-1_{\{l_m\in c_i-\}}]=x_m,1\leqslant m\leqslant s}\mathbb{P}\left(\tilde{J}_n^{c_1}=y_1,\cdots,\tilde{J}_n^{c_r}=y_r\right)\\
&= \sum_{\sum_{i=1}^r y_i[1_{\{l_m\in c_i\}}-1_{\{l_m\in c_i-\}}]=x_m,1\leqslant m\leqslant s}\mathbb{P}\left(\tilde{J}_n^{c_1}=-y_1,\cdots,\tilde{J}_n^{c_r}=-y_r\right)e^{n\sum_{i=1}^r y_i\log\frac{\gamma^{c_i}}{\gamma^{c_i-}}}\\
&= \sum_{\sum_{i=1}^r y_i[1_{\{l_m\in c_i\}}-1_{\{l_m\in c_i-\}}]=x_m,1\leqslant m\leqslant s}\mathbb{P}\left(\tilde{J}_n^{c_1}=-y_1,\cdots,\tilde{J}_n^{c_r}=-y_r\right)e^{n\sum_{i=1}^s x_i\log\frac{\gamma^{c_{l_i}}}{\gamma^{c_{l_i}-}}}\\
&= \sum_{\sum_{i=1}^r y_i[1_{\{l_m\in c_i\}}-1_{\{l_m\in c_i-\}}]=-x_m,1\leqslant m\leqslant s}\mathbb{P}\left(\tilde{J}_n^{c_1}=y_1,\cdots,\tilde{J}_n^{c_r}=y_r\right)e^{n\sum_{i=1}^s x_i\log\frac{\gamma^{c_{l_i}}}{\gamma^{c_{l_i}-}}}\\
&= \mathbb{P}\left(\sum_{i=1}^r\tilde{J}_n^{c_i}\left[1_{\{l_1\in c_i\}}-1_{\{l_1\in c_i-\}}\right]=-x_1,\cdots,\sum_{i=1}^r\tilde{J}_n^{c_i}\left[1_{\{l_s\in c_i\}}-1_{\{l_s\in c_i-\}}\right]=-x_s\right)e^{n\sum_{i=1}^s x_i\log\frac{\gamma^{c_{l_i}}}{\gamma^{c_{l_i}-}}}\\
&= \mathbb{P}\left(\tilde{Q}_n^{c_{l_1}}=x_1,\cdots,\tilde{Q}_n^{c_{l_s}}=x_s\right)e^{n\sum_{i=1}^s x_i\log\frac{\gamma^{c_{l_i}}}{\gamma^{c_{l_i}-}}},
\end{aligned}
$$

where we use the fact that under the constraint of $\sum_{i=1}^r y_i[1_{\{l_m\in c_i\}}-1_{\{l_m\in c_i-\}}]=x_m,\ \forall\,1\leqslant m\leqslant s$, we have

$$\sum_{i=1}^r y_i\log\frac{\gamma^{c_i}}{\gamma^{c_i-}}=\sum_{j=1}^s x_j\log\frac{\gamma^{c_{l_j}}}{\gamma^{c_{l_j}-}}.\tag{E1}$$

This identity is highly nontrivial and we next prove it. For any cycle c, let L^c be a function on E defined by

$$L^c(i,j) = \begin{cases} 1, & \text{if } \langle i,j\rangle \in c \\ 0, & \text{otherwise.} \end{cases}$$

By the definition of the function H^{c_l} in Eq. (21), it can be proved that [53]

$$L^c = \sum_{l\notin T} L^c(l) H^{c_l}.$$

Let w be a function on E defined by

$$w(i,j) = \log\frac{p_{ij}}{p_{ji}}.$$

For any cycle $c = (i_1, i_2, \cdots, i_t)$, we have

$$\log\frac{\gamma^c}{\gamma^{c-}} = \sum_{k=1}^{t}\log\frac{p_{i_k,i_{k+1}}}{p_{i_{k+1},i_k}} = \langle w, L^c\rangle,$$

where $i_{t+1} = i_1$ and $\langle w, L^c\rangle = \sum_{\langle i,j\rangle\in E} w(i,j)L^c(i,j)$ is the inner product. Moreover, for any $c_l \in \mathcal{L}$, it is not difficult to prove that

$$\log\frac{\gamma^{c_l}}{\gamma^{c_l-}} = \langle w, H^{c_l}\rangle.$$

Note that $\log(\gamma^c/\gamma^{c-}) = 0$ for all one-state or two-state cycles. Then for any cycle c, we have

$$\sum_{j=1}^{s}[L^c(l_j) - L^{c-}(l_j)]\log\frac{\gamma^{c_{l_j}}}{\gamma^{c_{l_j}-}} = \sum_{l\notin T} L^c(l)\log\frac{\gamma^{c_l}}{\gamma^{c_l-}}.$$

Thus, we finally obtain

$$\begin{aligned}\sum_{j=1}^{s} x_j\log\frac{\gamma^{c_{l_j}}}{\gamma^{c_{l_j}-}} &= \sum_{j=1}^{s}\sum_{i=1}^{r} y_i[L^{c_i}(l_j) - L^{c_i-}(l_j)]\log\frac{\gamma^{c_{l_j}}}{\gamma^{c_{l_j}-}} = \sum_{i=1}^{r} y_i\sum_{j=1}^{s}[L^{c_i}(l_j) - L^{c_i-}(l_j)]\log\frac{\gamma^{c_{l_j}}}{\gamma^{c_{l_j}-}} \\ &= \sum_{i=1}^{r} y_i\sum_{l\notin T} L^{c_i}(l)\langle w, H^{c_l}\rangle = \sum_{i=1}^{r} y_i\left\langle w, \sum_{l\notin T} L^{c_i}(l)H^{c_l}\right\rangle = \sum_{i=1}^{r} y_i\langle w, L^{c_i}\rangle = \sum_{i=1}^{r} y_i\log\frac{\gamma^{c_i}}{\gamma^{c_i-}}.\end{aligned}$$

This completes the proof of Eq. (E1) and thus completes the proof of the transient fluctuation theorem.

[1] C. Jarzynski, Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale, Annu. Rev. Condens. Matter Phys. **2**, 329 (2011).

[2] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Rep. Prog. Phys. **75**, 126001 (2012).

[3] C. Van den Broeck and M. Esposito, Ensemble and trajectory thermodynamics: A brief introduction, Physica A **418**, 6 (2015).

[4] L. Onsager and S. Machlup, Fluctuations and irreversible processes, Phys. Rev. **91**, 1505 (1953).

[5] L. Hong, C. Jia, Y. Zhu, and W.-A. Yong, Novel dissipative properties of the master equation, J. Math. Phys. **57**, 103303 (2016).

[6] A. Kolmogoroff, Zur theorie der Markoffschen ketten, Math. Ann. **112**, 155 (1936).

[7] C. Jia, D.-Q. Jiang, and Y. Li, Detailed balance, local detailed balance, and global potential for stochastic chemical reaction networks, Adv. Appl. Probab. **53**, 886 (2021).

[8] J. Schnakenberg, Network theory of microscopic and macroscopic behavior of master equation systems, Rev. Mod. Phys. **48**, 571 (1976).

[9] X.-J. Zhang, H. Qian, and M. Qian, Stochastic theory of nonequilibrium steady states and its applications. Part I, Phys. Rep. **510**, 1 (2012).

[10] H. Ge, M. Qian, and H. Qian, Stochastic theory of nonequilibrium steady states. Part II: Applications in chemical biophysics, Phys. Rep. **510**, 87 (2012).

[11] T. L. Hill and O. Kedem, Studies in irreversible thermodynamics III. Models for steady state and active transport across membranes, J. Theor. Biol. **10**, 399 (1966).

[12] T. L. Hill, Studies in irreversible thermodynamics IV. Diagrammatic representation of steady state fluxes for unimolecular systems, J. Theor. Biol. **10**, 442 (1966).

[13] T. L. Hill, Free energy and the kinetics of biochemical diagrams, including active transport, Biochemistry **14**, 2127 (1975).

[14] T. L. Hill, *Free Energy Transduction and Biochemical Cycle Kinetics* (Springer, New York, 1989).

[15] M. Qian and M. Qian, Circulation for recurrent Markov chains, Z. Wahrscheinlichkeitstheorie verw. Gebiete **59**, 203 (1982).

[16] M. Qian, M. Qian, and C. Qian, Circulations of Markov chains with continuous time and the probability interpretation of some determinants. Sci. Sin. Ser. A **27**, 470 (1984).

[17] D.-Q. Jiang, M. Qian, and M.-P. Qian, *Mathematical Theory of Nonequilibrium Steady States: On the Frontier of Probability and Dynamical Systems* (Springer, Berlin Heidelberg, 2004).

[18] E. Roldán and P. Vivo, Exact distributions of currents and frenesy for Markov bridges, Phys. Rev. E **100**, 042108 (2019).

[19] J. W. Biddle and J. Gunawardena, Reversal symmetries for cyclic paths away from thermodynamic equilibrium, Phys. Rev. E **101**, 062125 (2020).

[20] P. Pietzonka, J. Guioth, and R. L. Jack, Cycle counts and affinities in stochastic models of nonequilibrium systems, Phys. Rev. E **104**, 064137 (2021).

[21] D. J. Evans, E. G. Cohen, and G. P. Morriss, Probability of Second Law Violations in Shearing Steady States, Phys. Rev. Lett. **71**, 2401 (1993).

[22] G. Gallavotti, and E. G. D. Cohen, Dynamical ensembles in stationary states, J. Stat. Phys. **80**, 931 (1995).

[23] C. Jarzynski, Nonequilibrium Equality for Free Energy Differences, Phys. Rev. Lett. **78**, 2690 (1997).

[24] J. Kurchan, Fluctuation theorem for stochastic dynamics, J. Phys. A: Math. Gen. **31**, 3719 (1998).

[25] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, Phys. Rev. E **60**, 2721 (1999).

[26] D. J. Searles and D. J. Evans, Fluctuation theorem for stochastic systems, Phys. Rev. E **60**, 159 (1999).

[27] J. L. Lebowitz and H. Spohn, A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics, J. Stat. Phys. **95**, 333 (1999).

[28] T. Hatano and S.-I. Sasa, Steady-State Thermodynamics of Langevin Systems, Phys. Rev. Lett. **86**, 3463 (2001).

[29] U. Seifert, Entropy Production Along a Stochastic Trajectory and an Integral Fluctuation Theorem, Phys. Rev. Lett. **95**, 040602 (2005).

[30] M. Esposito and C. Van den Broeck, Three Detailed Fluctuation Theorems, Phys. Rev. Lett. **104**, 090601 (2010).

[31] H. K. Lee, C. Kwon, and H. Park, Fluctuation Theorems and Entropy Production with Odd-Parity Variables, Phys. Rev. Lett. **110**, 050602 (2013).

[32] H. Ge, C. Jia, and X. Jin, Martingale structure for general thermodynamic functionals of diffusion processes under second-order averaging, J. Stat. Phys. **184**, 17 (2021).

[33] D. Andrieux and P. Gaspard, Fluctuation theorem for currents and Schnakenberg network theory, J. Stat. Phys. **127**, 107 (2007).

[34] M. Polettini and M. Esposito, Transient fluctuation theorems for the currents and initial equilibrium ensembles, J. Stat. Mech.: Theory Exp. (2014) P10033.

[35] D. Andrieux and P. Gaspard, Network and thermodynamic conditions for a single macroscopic current fluctuation theorem, C. R. Phys. **8**, 579 (2007).

[36] C. Jia, D.-Q. Jiang, and M.-P. Qian, Cycle symmetries and circulation fluctuations for discrete-time and continuous-time Markov chains, Ann. Appl. Probab. **26**, 2454 (2016).

[37] H. Ge, C. Jia, and D.-Q. Jiang, Cycle symmetry, limit theorems, and fluctuation theorems for diffusion processes on the circle, Stoch. Proc. Appl. **127**, 1897 (2017).

[38] S. S. Varadhan, *Large Deviations and Applications* (Society for Industrial and Applied Mathematics, Philadelphia, 1984).

[39] F. Den Hollander, *Large Deviations* (American Mathematical Society, 2008), Vol. 14.

[40] L. Bertini, A. Faggionato, and D. Gabrielli, Flows, currents, and cycles for Markov chains: Large deviation asymptotics, Stoch. Proc. Appl. **125**, 2786 (2015).

[41] L. Bertini, A. Faggionato, and D. Gabrielli, Large deviations of the empirical flow for continuous time Markov chains, Ann. Inst. Henri Poincare, B **51**, 867 (2015).

[42] A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, 4th ed. (Wiley-Blackwell, Weinheim, 2012).

[43] B. Sakmann, *Single-Channel Recording*, 2nd ed. (Springer, US, 1995).

[44] C. Jia and R. Grima, Frequency Domain Analysis of Fluctuations of mRNA and Protein Copy Numbers within a Cell Lineage: Theory and Experimental Validation, Phys. Rev. X **11**, 021032 (2021).

[45] C. Jia, A. Singh, and R. Grima, Cell size distribution of lineage data: Analytic results and parameter inference, iScience **24**, 102220 (2021).

[46] C. Jia, A. Singh, and R. Grima, Characterizing non-exponential growth and bimodal cell size distributions in fission yeast: An analytical approach, PLOS Computational Biology **18(1)**, e1009793 (2022).

[47] P. B. Gupta *et al.*, Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells, Cell **146**, 633 (2011).

[48] D. Zhou, D. Wu, Z. Li, M. Qian, and M. Q. Zhang, Population dynamics of cancer cells with cell state conversions, Quantum Biol. **1**, 201 (2013).

[49] H. Qian, Phosphorylation energy hypothesis: Open chemical systems and their biological functions, Annu. Rev. Phys. Chem. **58**, 113 (2007).

[50] D. A. Beard and H. Qian, *Chemical Biophysics: Quantitative Analysis of Cellular Systems* (Cambridge University Press, Cambridge, 2008), Vol. 126.

[51] J. M. Pedraza and J. Paulsson, Effects of molecular memory and bursting on fluctuations in gene expression, Science **319**, 339 (2008).

[52] D. M. Suter *et al.*, Mammalian genes are transcribed with widely different bursting kinetics, Science **332**, 472 (2011).

[53] S. L. Kalpazidou, *Cycle Representations of Markov Processes* (Springer, New York, 2007), Vol. 28.

[54] H. Ge, Waiting cycle times and generalized Haldane equality in the steady-state cycle kinetics of single enzymes, J. Phys. Chem. B **112**, 61 (2008).

[55] J. R. Norris, J. R. Norris, and J. R. Norris, *Markov Chains*, 2nd ed. (Cambridge University Press, Cambridge, 1998).

[56] J. Paulsson, Models of stochastic gene expression, Phys. Life Rev. **2**, 157 (2005).

[57] C. Jia, A. Singh, and R. Grima, Concentration fluctuations in growing and dividing cells: Insights into the emergence of concentration homeostasis, PLOS Computational Biology **18(10)**, e1010574 (2022).

[58] J. M. Schmiedel, S. L. Klemm, Y. Zheng, A. Sahay, N. Blüthgen, D. S. Marks, and A. Van Oudenaarden, MicroRNA control of protein expression noise, Science **348**, 128 (2015).

[59] S. Bo and A. Celani, Multiple-scale stochastic processes: Decimation, averaging and beyond, Phys. Rep. **670**, 1 (2017).

[60] C. Jia, Model simplification and loss of irreversibility, Phys. Rev. E **93**, 052149 (2016).

[61] H. Qian and X. S. Xie, Generalized Haldane equation and fluctuation theorem in the steady-state cycle kinetics of single enzymes, Phys. Rev. E **74**, 010902(R) (2006).

[62] J. Mehl, B. Lander, C. Bechinger, V. Blickle, and U. Seifert, Role of Hidden Slow Degrees of Freedom in the Fluctuation Theorem, Phys. Rev. Lett. **108**, 220601 (2012).

[63] M. Polettini and M. Esposito, Effective Thermodynamics for a Marginal Observer, Phys. Rev. Lett. **119**, 240601 (2017).

[64] M. Uhl, P. Pietzonka, and U. Seifert, Fluctuations of apparent entropy production in networks with hidden slow degrees of freedom, J. Stat. Mech.: Theory Exp. (2018) 023203.

[65] M. Kahlen and J. Ehrich, Hidden slow degrees of freedom and fluctuation theorems: An analytically solvable model, J. Stat. Mech.: Theory Exp. (2018) 063204.