The Annals of Applied Probability
2016, Vol. 26, No. 4, 2454–2493
DOI: 10.1214/15-AAP1152

CYCLE SYMMETRIES AND CIRCULATION FLUCTUATIONS FOR DISCRETE-TIME AND CONTINUOUS-TIME MARKOV CHAINS[1]

BY CHEN JIA*,†, DA-QUAN JIANG† AND MIN-PING QIAN†

Beijing Computational Science Research Center and Peking University†*

In this paper, we find a series of equalities which characterize the symmetry of the forming times of a family of similar cycles for discrete-time and continuous-time Markov chains. Moreover, we use these cycle symmetries to study the circulation fluctuations for Markov chains. We prove that the sample circulations along a family of cycles passing through a common state satisfy a large deviation principle with a rate function which has a highly nonobvious symmetry. Further extensions and applications to statistical physics and biochemistry are also discussed, especially the fluctuation theorems for the sample net circulations.

1. Introduction. Markov chains are widely used to model various stochastic systems in physics, chemistry, biology, engineering and other disciplines. The trajectory of a recurrent Markov chain constantly forms various cycles. The cycle representation theory of Markov chains [23, 31, 34–38] not only possesses rich theoretical results, but has become a fundamental tool in dealing with nonequilibrium (irreversible) processes in natural sciences as well. Readers may refer to [21, 24] for the theoretical contents of the cycle representation theory and refer to [18, 45] for its applications in physics, chemistry and biology.

The earliest theoretical result about the cycle representation theory is probably Kolmogorov's criterion for reversibility [25], which claims that a stationary Markov chain is reversible if and only if the product of transition probabilities (rates) along each cycle c and that along its reversed cycle $c-$ are the same. Illuminated by Hill's diagram method [19] and Schnakenberg's network theory [40], the Qian's [31, 34, 35, 37, 38] and Kalpazidou [23, 24] introduced the important concept of circulations and further enriched the cycle representation theory. Let N_t^c denote the number of times that cycle c is formed by a Markov chain up to time t. Then the sample circulation J_t^c along cycle c by time t is defined as

$$J_t^c = \frac{1}{t} N_t^c \tag{1}$$

Received July 2014; revised September 2015.

[1]Supported by NSFC (Nos 11271029 and 11171024).

MSC2010 subject classifications. 60J10, 60J20, 60J27, 60J28, 60F10, 82C31.

Key words and phrases. Haldane equality, nonequilibrium, fluctuation theorem, current fluctuation, large deviation.

and the circulation J^c along cycle c is a nonnegative real number defined as the following almost sure limit:

$$J^c = \lim_{t\to\infty} J_t^c, \qquad \text{a.s.,} \tag{2}$$

which represents the number of times that cycle c is formed per unit time. It turns out that a stationary Markov chain is reversible if and only if the circulations along each cycle c and its reversed cycle $c-$ are the same. This explains why the cycle representation theory is naturally related to various nonequilibrium phenomena in natural sciences.

Recently, biophysicists have applied the cycle representation theory to study single-molecule enzyme kinetics and found an interesting relation named as the generalized Haldane equality [16–18, 33]. Mathematically, each chemical reaction catalyzed by an enzyme can be modeled as a Markov chain whose state space has a cyclic topology (see Section 7.2 below). For such cyclic Markov chains, there are only two effective cycles, that is, the clockwise cycle c and the counterclockwise cycle $c-$. Let T^c and T^{c-} denote the time needed for the Markov chain to form the cycle c and its reversed cycle $c-$ for the first time, respectively. Kolomeisky et al. [26] proved that for such cyclic Markov chains, the expectations of T^c and T^{c-}, under the condition that the corresponding cycle is formed earlier than its reversed cycle, are exactly the same:

$$E\left(T^c|T^c < T^{c-}\right) = E\left(T^{c-}|T^{c-} < T^c\right). \tag{3}$$

Subsequently, Qian and Xie [33] and Ge [16] generalized the equality (3) and proved that for cyclic Markov chains, not only the conditional expectations, but also the conditional distributions of T^c and T^{c-} are also the same:

$$P\left(T^c \le t|T^c < T^{c-}\right) = P\left(T^{c-} \le t|T^{c-} < T^c\right). \tag{4}$$

This equality characterizes a symmetry of the forming times of a cycle and its reversed cycle. Qian and Xie [33] named the equality (4) as the generalized Haldane equality since it generalizes what is known as the Haldane relation for reversible enzyme kinetics [41]. Interestingly, Samuels [39] and Dubins [10] have found an equivalent form of the generalized Haldane equality even earlier in nearest-neighbor periodic walks when studying the gambler's ruin problem.

Now that the generalized Haldane equality holds for cyclic Markov chains, it is natural to ask whether it holds for general Markov chains. If the state space has a cyclic topology, then the Markov chain has only two effective cycles. In this case, the generalized Haldane equality can be proved by the method of quasi-time-reversal [16]. However, this method depends too much on the cyclic topology of the state space and cannot be generalized to general Markov chains with a large number of effective cycles.

In this paper, we establish some deep properties of taboo probabilities and use them to prove the generalized Haldane equality for general discrete-time and continuous-time Markov chains with denumerable state space. We find that the generalized Haldane equality not only holds for a cycle and its reversed cycle, but also holds for a family of similar cycles, which are defined as cycles passing through the same set of states (see Definition 3.1 below). Let $c_1, c_2, \ldots, c_r$ be a family of similar cycles, let $T^{c_1}, T^{c_2}, \ldots, T^{c_r}$ be their forming times and let $T = \min\{T^{c_1}, T^{c_2}, \ldots, T^{c_r}\}$. In this paper, we prove that although the distributions of $T^{c_1}, T^{c_2}, \ldots, T^{c_r}$ can be different, their distributions, under the condition that the corresponding cycle is formed earlier than other similar cycles, are exactly the same:

$$P\left(T^{c_1} \leq t | T = T^{c_1}\right) = P\left(T^{c_2} \leq t | T = T^{c_2}\right) = \cdots = P\left(T^{c_r} \leq t | T = T^{c_r}\right). \tag{5}$$

This equality also implies that the forming time T of multiple similar cycles is independent of which one of these cycles is formed (see Remark 3.13 below). This is another important aspect of the generalized Haldane equality.

The generalized Haldane equality established in this paper has wide applications. One of the most important applications is to study the circulation fluctuations for Markov chains. In the recent two decades, the studies on the fluctuations of various thermodynamic quantities for stochastic systems have become a central topic in nonequilibrium statistical physics [43]. Motivated by the results of numerical simulations [13], Gallavotti and Cohen [15] gave the first mathematical presentation of the fluctuation theorem for a class of stationary nonequilibrium systems. Since then, there has been a large amount of literature exploring various generalizations of the fluctuation theorem [8, 12, 20, 27, 28, 42, 44].

In recent years, physicists became increasingly concerned about the fluctuation theorems for the circulations of Markov chains [1–4, 14, 18, 30, 33]. As pointed out by Seifert [43], the studies on the circulation fluctuations have become a hot spot topic in nonequilibrium statistical physics. In the previous work, physicists have studied the fluctuation theorems for the cycle currents of Markov chains based on Schnakenberg's network theory [1, 14, 30]. However, the cycle currents studied by physicists are related to but in essence different from the circulations considered in this paper [14]. In addition, Andrieux and Gaspard [2] have tried to study the circulation fluctuations for Markov chains. However, their proofs of the fluctuation theorems are not mathematically rigorous and seem questionable in some complicated cases. As a result, the circulation fluctuations for Markov chains need to be studied in more detail with full mathematical rigor.

Interestingly, the generalized Haldane equality established in this paper can be used to study the circulation fluctuations for Markov chains. It is easy to see that the circulations defined in (2) are the almost sure limits of the sample circulations defined in (1). In this paper, we prove that the sample circulations along a family of cycles $c_1, c_2, \ldots, c_r$ passing through a common state satisfy a large deviation principle with rate t and good rate function $I^{c_1,c_2,\ldots,c_r}$. Moreover, we apply the generalized Haldane equality to prove that the rate function $I^{c_1,c_2,\dots,c_r}$ has the following highly nonobvious symmetry: if c_k and c_l are similar, then

$$
\begin{aligned}
&I^{c_1,c_2,\dots,c_r}(x_1,\dots,x_k,\dots,x_l,\dots,x_r) \\
&\quad = I^{c_1,c_2,\dots,c_r}(x_1,\dots,x_l,\dots,x_k,\dots,x_r) - \left(\log\frac{\gamma^{c_k}}{\gamma^{c_l}}\right)(x_k - x_l),
\end{aligned}
\tag{6}
$$

where γ^{c_k} and γ^{c_l} are the strengths of c_k and c_l, respectively (see Definition 5.1 below).

The results of this paper can be directly applied to statistical physics. In nonequilibrium statistical physics, one of the topics of interest is the fluctuation theorems for the sample net circulations, where the sample net circulation K_t^c along cycle c by time t is defined as

$$
K_t^c = J_t^c - J_t^{c-}. \tag{7}
$$

In this paper, we prove that the sample net circulations along cycles $c_1, c_2, \dots, c_r$ also satisfy a large deviation principle with rate t and good rate function $I_K^{c_1,c_2,\dots,c_r}$ which has the following symmetry:

$$
\begin{aligned}
&I_K^{c_1,c_2,\dots,c_r}(x_1,\dots,x_k,\dots,x_r) \\
&\quad = I^{c_1,c_2,\dots,c_r}(x_1,\dots,-x_k,\dots,x_r) - \left(\log\frac{\gamma^{c_k}}{\gamma^{c_k-}}\right)x_k.
\end{aligned}
\tag{8}
$$

This is actually the Gallavotti–Cohen-type fluctuation theorem for the sample net circulations. During the proof of this result, we also obtain other types of fluctuation theorems as by-products, including the transient fluctuation theorem, the integral fluctuation theorem and the Kurchan–Lebowitz–Spohn-type fluctuation theorem. All these fluctuation theorems, together with the generalized Haldane equality, characterize the symmetries of Markov chains along a family of similar cycles from different aspects.

At the end of this paper, we briefly discuss the application of our work to biochemistry. This indicates that our work could have a broad application prospect in natural sciences.

2. Rigorous definitions of cycles and their forming times.

In this section, we shall give the rigorous definitions of cycles and their forming times for discrete-time and continuous-time Markov chains.

We first give the definition of cycles. Here, we adopt the definition given by Kalpazidou [24]. Let $X = (X_t)_{t\geq 0}$ be a time-homogeneous discrete-time or continuous-time Markov chain with denumerable state space S defined on some probability space $(\Omega, \mathscr{F}, P)$.

DEFINITION 2.1. Let $\mathbb{Z}$ be the set of integers. A circuit function in the state space S is defined as a periodic function f from $\mathbb{Z}$ into S. The smallest positive integer s such that $f(n+s) = f(n)$ for each $n \in \mathbb{Z}$ is called the period of f.

DEFINITION 2.2. Two circuit functions f and g in S are called equivalent if there exists some $k \in \mathbb{Z}$ such that $g(n) = f(n+k)$ for each $n \in \mathbb{Z}$.

Note that Definition 2.2 introduces an equivalence relation on the space of all circuit functions in S. It is obvious that two equivalent circuit functions have the same period.

DEFINITION 2.3. Let $i_1, i_2, \dots, i_s$ be distinct states in S. Let f be a circuit function in S with period s that satisfies $f(1) = i_1$, $f(2) = i_2, \dots, f(s) = i_s$. Then the equivalence class of the circuit function f under the equivalence relation described in Definition 2.2 is called a cycle and is denoted by $(i_1, i_2, \dots, i_s)$.

In other words, a cycle is nothing but an equivalence class on the space of all circuit functions under the equivalence relation described in Definition 2.2. According to the above definition, two cycles are the same if one can be transformed into the other by a cyclic permutation. For example, (1, 2, 3), (2, 3, 1) and (3, 1, 2) represent the same cycle.

DEFINITION 2.4. Let $i \in S$ and let $c = (i_1, i_2, \dots, i_s)$ be a cycle. Then we say that cycle c passes through state i if $i \in \{i_1, i_2, \dots, i_s\}$.

We next give the definition of the forming times of cycles for discrete-time Markov chains. To this end, we must first introduce the definition of the derived chain. Let $X = (X_n)_{n \ge 0}$ be an irreducible and recurrent discrete-time Markov chain with denumerable state space S and transition probability matrix $P = (p_{ij})$.

The trajectory of a recurrent Markov chain constantly forms various cycles. Intuitively, if we discard the cycles formed by X and keep track of the remaining states in the trajectory, then we obtain a new Markov chain Y called the derived chain. We shall give the rigorous definition of the derived chain later, but the basic ideas should be clear from the following example.

EXAMPLE 2.5. If the trajectory of the Markov chain X is $\{1, 2, 3, 2, 4, 5, 2, 3, 1, \dots\}$, then the corresponding trajectory of the derived chain Y and the cycles formed are as follows (see Table 1).

TABLE 1
An example of the derived chain and the cycles formed

n	0	1	2	3	4	5	6	7	8
X_n	1	2	3	2	4	5	2	3	1
Y_n	[1]	[1, 2]	[1, 2, 3]	[1, 2]	[1, 2, 4]	[1, 2, 4, 5]	[1, 2]	[1, 2, 3]	[1]
Cycles formed				(2, 3)			(2, 4, 5)		(1, 2, 3)

In order to give the rigorous definition of the derived chain, we introduce several notation. We denote an finite sequence $i_1, i_2, \ldots, i_s$ of distinct states by $[i_1, i_2, \ldots, i_s]$ and denote the collection of all finite sequences of distinct states by $[S]$, that is,

$$[S] = \{[i_1, i_2, \ldots, i_s] : s \geq 1, i_1, \ldots, i_s \text{ are distinct states in } S\}. \tag{9}$$

It is easy to see that $[S]$ is denumerable. We also define a map $\{\cdot, \cdot\}$ from $[S] \times S$ into $[S]$ by

$$\begin{aligned} &\{[i_1, i_2, \ldots, i_s], i\} \\ &\quad = \begin{cases} [i_1, i_2, \ldots, i_s, i], & \text{if } i \notin \{i_1, i_2, \ldots, i_s\}, \\ [i_1, i_2, \ldots, i_m], & \text{if } i = i_m \text{ for some } 1 \leq m \leq s. \end{cases} \end{aligned} \tag{10}$$

DEFINITION 2.6. The derived chain $Y = (Y_n)_{n \geq 0}$ of X is defined as $Y_0 = [X_0]$ and $Y_n = \{Y_{n-1}, X_n\}$ for each $n \geq 1$.

For any $i \in S$, let $[S]_i$ be the subset of $[S]$ defined by

$$[S]_i = \{[i_1, i_2, \ldots, i_s] \in [S] : i_1 = i \text{ and } p_{i_1 i_2} p_{i_2 i_3} \cdots p_{i_{s-1} i_s} > 0\}. \tag{11}$$

It is easy to see that if $Y_0 = [i]$, then $Y_n \in [S]_i$ for each $n \geq 1$.

PROPOSITION 2.7. *The derived chain Y is a time-homogeneous Markov chain with denumerable state space $[S]$. Each $[S]_i$ is an irreducible recurrent class of Y.*

PROOF. Let $y_1 = [i_1, i_2, \ldots, i_s]$ and $y_2 = [j_1, j_2, \ldots, j_r]$ be two states in $[S]$. It is easy to see that Y is a time-homogeneous Markov chain on $[S]$ with transition probability

$$p_{y_1 y_2} = \begin{cases} p_{i_s j_r}, & \text{if } r \leq s \text{ and } i_1 = j_1, i_2 = j_2, \ldots, i_r = j_r, \\ p_{i_s j_r}, & \text{if } r = s + 1 \text{ and } i_1 = j_1, i_2 = j_2, \ldots, i_s = j_s, \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

Moreover, it is easy to see that $[S]_i$ is a communicating class of Y. If $Y_0 = [i]$, then $Y_n = [i]$ if and only if $X_n = i$ for each $n \geq 0$. Since X is recurrent, X will return to i infinitely often, which means that Y will return to $[i]$ infinitely often. Thus, $[i]$ is a recurrent state of Y. Since recurrence is a property of the communicating class, we claim that $[S]_i$ is an irreducible recurrent class of Y. $\square$

DEFINITION 2.8. Let $c = (i_1, i_2, \ldots, i_s)$ be a cycle. For each $\omega \in \Omega$, we say that the trajectory $X(\omega)$ forms cycle c at time n if one of the following two cases occurs:

(i) there exists $1 \leq m \leq s$ such that $Y_{n-1}(\omega) = [i_m, i_{m+1}, \ldots, i_{m+s-1}]$ and $Y_n(\omega) = [i_m]$;

(ii) there exists $1 \le m \le s$ and distinct states $j_1, j_2, \dots, j_r \notin \{i_1, i_2, \dots, i_s\}$ such that $Y_{n-1}(\omega) = [j_1, j_2, \dots, j_r, i_m, i_{m+1}, \dots, i_{m+s-1}]$ and $Y_n(\omega) = [j_1, j_2, \dots, j_r, i_m]$,

where $m+1, m+2, \dots, m+s-1$ are understood to be modulo s.

DEFINITION 2.9. Let c be a cycle. Let $(T_n^c)_{n\ge1}$ be a sequence of stopping times defined by

$$\begin{aligned} T_1^c(\omega) &= \inf\{k \ge 1 : \text{the trajectory } X(\omega) \text{ forms cycle } c \text{ at time } k\}, \\ T_n^c(\omega) &= \inf\{k \ge T_{n-1}^c(\omega) + 1 : \text{the trajectory } X(\omega) \\ &\qquad \text{forms cycle } c \text{ at time } k\} \qquad \forall n \ge 2. \end{aligned} \tag{13}$$

Then T_n^c is called the nth forming time of cycle c. The (first) forming time T_1^c of cycle c is always abbreviated as T^c in the following discussion.

PROPOSITION 2.10. *Let $c = (i_1, i_2, \dots, i_s)$ be a cycle.*

(i) *If $p_{i_1i_2}p_{i_2i_3}\cdots p_{i_si_1} > 0$, then $T_n^c < \infty$, a.s. for each $n \ge 1$.*
(ii) *If $p_{i_1i_2}p_{i_2i_3}\cdots p_{i_si_1} = 0$, then $T_n^c = \infty$, a.s. for each $n \ge 1$.*

PROOF. It is easy to see that (ii) holds. We next prove (i). Without loss of generality, we assume that X starts from an arbitrarily fixed state i.

We first consider the case when c passes through i. Since $p_{i_1i_2}p_{i_2i_3}\cdots p_{i_si_1} > 0$, there exists $1 \le m \le s$ such that $i_m = i$ and $y = [i_m, i_{m+1}, \dots, i_{m+s-1}] \in [S]_i$. By Proposition 2.7, the derived chain Y will hit y infinitely often. Since $p_{i_{m-1}i_m} > 0$, by the Markov property of Y, there will be infinitely many n such that $Y_{n-1} = y = [i_m, i_{m+1}, \dots, i_{m+s-1}]$ and $Y_n = [i_m]$. This shows that $T_n^c < \infty$, a.s. for each $n \ge 1$.

We next consider the case when c does not pass through i. Since $p_{i_1i_2}p_{i_2i_3}\cdots p_{i_si_1} > 0$, there exists $1 \le m \le s$ and distinct states $j_1, j_2, \dots, j_r \notin \{i_1, i_2, \dots, i_s\}$ such that $j_1 = i$ and $z = [j_1, j_2, \dots, j_r, i_m, i_{m+1}, \dots, i_{m+s-1}] \in [S]_i$. By Proposition 2.7, the derived chain Y will hit z infinitely often. Since $p_{i_{m-1}i_m} > 0$, by the Markov property of Y, there will be infinitely many n such that $Y_{n-1} = z = [j_1, j_2, \dots, j_r, i_m, i_{m+1}, \dots, i_{m+s-1}]$ and $Y_n = [j_1, j_2, \dots, j_r, i_m]$. This shows that $T_n^c < \infty$, a.s. for each $n \ge 1$. $\square$

We finally give the definition of the forming times of cycles for continuous-time Markov chains. Let $X = (X_t)_{t\ge0}$ be an irreducible and recurrent continuous-time Markov chain with denumerable state space S and transition rate matrix $Q = (q_{ij})$. Let $(J_n)_{n\ge0}$ be the jump times of X, where J_0 is understood to be 0. For each $n \ge 0$, let $\bar{X}_n = X_{J_n}$. Then $\bar{X} = (\bar{X}_n)_{n\ge0}$ is the embedded chain of X.

DEFINITION 2.11. Let c be a cycle. For each $n \geq 1$, let $\bar{T}_n^c$ be the nth forming time of cycle c by the embedded chain $\bar{X}$. Then the nth forming time of cycle c by X is defined as

$$T_n^c = J_{\bar{T}_n^c}. \tag{14}$$

The (first) forming time T_1^c of cycle c is always abbreviated as T^c in the following discussion.

PROPOSITION 2.12. *Let $c = (i_1, i_2, \ldots, i_s)$ be a cycle.*

(i) *If $q_{i_1 i_2} q_{i_2 i_3} \cdots q_{i_s i_1} > 0$, then $T_n^c < \infty$, a.s. for each $n \geq 1$.*
(ii) *If $q_{i_1 i_2} q_{i_2 i_3} \cdots q_{i_s i_1} = 0$, then $T_n^c = \infty$, a.s. for each $n \geq 1$.*

PROOF. It is easy to see that (ii) holds. We next prove (i). Since X is irreducible and recurrent, the embedded chain $\bar{X}$ is also irreducible and recurrent. By Proposition 2.10, we see that $\bar{T}_n^c < \infty$, a.s. This shows that $T_n^c = J_{\bar{T}_n^c} < \infty$, a.s. $\square$

3. Generalized Haldane equality for discrete-time Markov chains.

In this section, we shall prove the generalized Haldane equality for discrete-time Markov chains. Let $X = (X_n)_{n \geq 0}$ be an irreducible and recurrent discrete-time Markov chain with denumerable state space S and transition probability matrix $P = (p_{ij})$.

Before we state the generalized Haldane equality, we introduce the following definition.

DEFINITION 3.1. Let $c_1 = (i_1, i_2, \ldots, i_s)$ and $c_2 = (j_1, j_2, \ldots, j_r)$ be two cycles. Then c_1 and c_2 are called similar if $s = r$ and $\{i_1, i_2, \ldots, i_s\} = \{j_1, j_2, \ldots, j_r\}$.

According to the above definition, two cycles are similar if they pass through the same set of states. It is easy to see that similarity is an equivalence relation on the space of all cycles. For example, the six cycles, $c_1 = (1, 2, 3, 4)$, $c_2 = (1, 2, 4, 3)$, $c_3 = (1, 3, 2, 4)$, $c_4 = (1, 3, 4, 2)$, $c_5 = (1, 4, 2, 3)$ and $c_6 = (1, 4, 3, 2)$, are similar.

We next give the definition of the strengths of cycles for discrete-time Markov chains.

DEFINITION 3.2. Let $c = (i_1, i_2, \ldots, i_s)$ be a cycle. Then the strength of cycle c is defined as

$$\gamma^c = p_{i_1 i_2} p_{i_2 i_3} \cdots p_{i_s i_1}. \tag{15}$$

The generalized Haldane equality, which characterizes the symmetry of the forming times of a family of similar cycles, is stated in the following theorem.

THEOREM 3.3. *Let* $c_1, c_2, \dots, c_r$ *be a family of similar cycles. Let* $T = \min\{T^{c_1}, T^{c_2}, \dots, T^{c_r}\}$. *Then*:

(i) *for each* $n \geq 1$ *and any* $1 \leq k, l \leq r$,

$$\frac{P(T^{c_k} = n, T = T^{c_k})}{P(T^{c_l} = n, T = T^{c_l})} = \frac{P(T = T^{c_k})}{P(T = T^{c_l})} = \frac{\gamma^{c_k}}{\gamma^{c_l}}; \tag{16}$$

(ii) *for each* $n \geq 1$,

$$\begin{aligned} P(T^{c_1} = n | T = T^{c_1}) &= P(T^{c_2} = n | T = T^{c_2}) \\ &= \cdots = P(T^{c_r} = n | T = T^{c_r}). \end{aligned} \tag{17}$$

REMARK 3.4. The above theorem shows that although the distributions of the forming times of a family of similar cycles may not be the same, their distributions, under the condition that the corresponding cycle is formed earlier than other similar cycles, are exactly the same. This is the first aspect of the generalized Haldane equality.

REMARK 3.5. It may occur that both the numerator and denominator in (16) are 0. In this case, (16) is understood to hold trivially. In addition, if $P(T = T^{c_k}) = 0$ for some k, then (17) is understood to hold trivially. This understanding applies to similar equalities below.

In order to prove the generalized Haldane equality, we need to establish some nontrivial properties of taboo probabilities. Let us first recall the definition of taboo probabilities, also called transition probabilities with a taboo set [7].

DEFINITION 3.6. Let $i, j \in S$ and let H be a subset of S. Then the n-step transition probability from state i to state j with taboo set H is defined as

$$p_{ij}^{H}(n) = P_i(X_n = j, X_1, \dots, X_{n-1} \notin H), \tag{18}$$

where $P_i(\cdot) = P(\cdot | X_0 = i)$. If the taboo set is the union of a set H and a finite number of states $k_1, \dots, k_s$, then we shall denote the taboo probability by $p_{ij}^{H,k_1,\dots,k_s}(n)$, that is,

$$p_{ij}^{H,k_1,\dots,k_s}(n) = P_i(X_n = j, X_1, \dots, X_{n-1} \notin H \cup \{k_1, \dots, k_s\}). \tag{19}$$

The next four lemmas give some deep properties of taboo probabilities. The following lemma is called the basic decomposition formula of taboo probabilities ([7], Theorem 1, Section 9, Part I). To make the paper self-contained, we give a proof of this lemma.

LEMMA 3.7. *Let H be a subset of S and let $k \notin H$. Then for each $n \geq 0$ and any $i, j \in S$,*

$$p_{ij}^{H}(n) = p_{ij}^{H,k}(n) + \sum_{m=1}^{n-1} p_{ik}^{H}(m) p_{kj}^{H,k}(n-m). \tag{20}$$

PROOF. When $n = 0$ or $n = 1$, it is easy to check that the theorem holds. We next prove the theorem when $n \geq 2$. Note that

$$p_{ij}^{H}(n) = p_{ij}^{H,k}(n) + P_i\big(X_n = j, X_1, \ldots, X_{n-1} \notin H, k \in \{X_1, \ldots, X_{n-1}\}\big). \tag{21}$$

Then by the Markov property, we obtain that

$$\begin{aligned}
&P_i\big(X_n = j, X_1, \ldots, X_{n-1} \notin H, k \in \{X_1, \ldots, X_{n-1}\}\big)\\
&\quad = \sum_{m=1}^{n-1} P_i(X_n = j, X_1, \ldots, X_{n-1} \notin H, X_m = k, X_{m+1}, \ldots, X_{n-1} \neq k)\\
&\quad = \sum_{m=1}^{n-1} P_i(X_m = k, X_1, \ldots, X_{m-1} \notin H)\\
&\qquad \times P_k\big(X_{n-m} = j, X_1, \ldots, X_{n-m-1} \notin H \cup \{k\}\big)\\
&\quad = \sum_{m=1}^{n-1} p_{ik}^{H}(m) p_{kj}^{H,k}(n-m).
\end{aligned}$$

This completes the proof of this lemma. □

LEMMA 3.8. *Let H be a subset of S. Let $i, j \notin H$ and $i \neq j$. Then for each $n \geq 0$,*

$$\sum_{m=0}^{n} p_{ii}^{H}(m) p_{jj}^{H,i}(n-m) = \sum_{m=0}^{n} p_{jj}^{H}(m) p_{ii}^{H,j}(n-m). \tag{22}$$

PROOF. By Lemma 3.7, we have

$$\begin{aligned}
&\sum_{m=0}^{n} p_{ii}^{H}(m) p_{jj}^{H,i}(n-m)\\
&\quad = \sum_{m=0}^{n} p_{ii}^{H}(m) p_{jj}^{H}(n-m) - \sum_{m=0}^{n} p_{ii}^{H}(m) \sum_{l=1}^{n-m-1} p_{ji}^{H}(l) p_{ij}^{H,i}(n-m-l)\\
&\quad = \sum_{m=0}^{n} p_{ii}^{H}(m) p_{jj}^{H}(n-m) - \sum_{m=0}^{n} p_{ii}^{H}(m) \sum_{l=0}^{n-m} p_{ji}^{H}(l) p_{ij}^{H,i}(n-m-l)\\
&\quad = \sum_{m=0}^{n} p_{ii}^{H}(m) p_{jj}^{H}(n-m) - \sum_{l=0}^{n} p_{ji}^{H}(l) \sum_{m=0}^{n-l} p_{ii}^{H}(m) p_{ij}^{H,i}(n-m-l).
\end{aligned} \tag{23}$$

Using Lemma 3.7 again, we have

$$
\begin{aligned}
&\sum_{m=0}^{n-l} p_{ii}^{H}(m)p_{ij}^{H,i}(n-m-l)\\
&\quad = p_{ij}^{H,i}(n-l) + \sum_{m=1}^{n-l-1} p_{ii}^{H}(m)p_{ij}^{H,i}(n-m-l) \qquad (24)\\
&\quad = p_{ij}^{H,i}(n-l) + p_{ij}^{H}(n-l) - p_{ij}^{H,i}(n-l) = p_{ij}^{H}(n-l).
\end{aligned}
$$

Thus, we obtain that

$$
\begin{aligned}
&\sum_{m=0}^{n} p_{ii}^{H}(m)p_{jj}^{H,i}(n-m)\\
&\qquad = \sum_{m=0}^{n} p_{ii}^{H}(m)p_{jj}^{H}(n-m) - \sum_{l=0}^{n} p_{ji}^{H}(l)p_{ij}^{H}(n-l). \qquad (25)
\end{aligned}
$$

Commuting i and j in the above equation, we finally obtain that

$$
\begin{aligned}
&\sum_{m=0}^{n} p_{jj}^{H}(m)p_{ii}^{H,j}(n-m)\\
&\qquad = \sum_{m=0}^{n} p_{jj}^{H}(m)p_{ii}^{H}(n-m) - \sum_{l=0}^{n} p_{ij}^{H}(l)p_{ji}^{H}(n-l)\\
&\qquad = \sum_{m=0}^{n} p_{ii}^{H}(m)p_{jj}^{H}(n-m) - \sum_{l=0}^{n} p_{ji}^{H}(l)p_{ij}^{H}(n-l) \qquad (26)\\
&\qquad = \sum_{m=0}^{n} p_{ii}^{H}(m)p_{jj}^{H,i}(n-m),
\end{aligned}
$$

which gives the desired result. $\square$

LEMMA 3.9. *Let H be a subset of S. For any distinct states $i_1, i_2, \dots, i_s \notin H$, let*

$$
G_n^H(i_1, i_2, \dots, i_s) = \sum_{n_1+n_2+\cdots+n_s=n} p_{i_1i_1}^{H}(n_1)p_{i_2i_2}^{H,i_1}(n_2)\cdots p_{i_si_s}^{H,i_1,\dots,i_{s-1}}(n_s). \qquad (27)
$$

Then for each $n \geq 0$, $G_n^H(i_1, i_2, \dots, i_s)$ is invariant under any permutation of $i_1, i_2, \dots, i_s$.

PROOF. Since any permutation can be decomposed into the product of some transpositions of adjacent elements, we only need to prove that $G_n^H(i_1, i_2, \dots, i_s)$

is invariant if we exchange two adjacent elements, i_k and i_{k+1}, and keep all other elements fixed. By Lemma 3.8, we obtain that

$$\begin{aligned}
&G_n^H(i_1,\dots,i_k,i_{k+1},\dots,i_s)\\
&\quad=\sum_{n_1+\cdots+n_s=n} p_{i_1i_1}^{H}(n_1)\cdots p_{i_ki_k}^{H,i_1,\dots,i_{k-1}}(n_k)\\
&\qquad\times p_{i_{k+1}i_{k+1}}^{H,i_1,\dots,i_k}(n_{k+1})\cdots p_{i_si_s}^{H,i_1,\dots,i_{s-1}}(n_s)\\
&\quad=\sum_{m=0}^{n}\sum_{n_1+\cdots+n_{k-1}+n_{k+2}+\cdots+n_s=n-m} p_{i_1i_1}^{H}(n_1)\cdots p_{i_{k-1}i_{k-1}}^{H,i_1,\dots,i_{k-2}}(n_{k-1})\\
&\qquad\times p_{i_{k+2}i_{k+2}}^{H,i_1,\dots,i_{k+1}}(n_{k+2})\cdots p_{i_si_s}^{H,i_1,\dots,i_{s-1}}(n_s)\\
&\qquad\times\sum_{n_k+n_{k+1}=m} p_{i_ki_k}^{H,i_1,\dots,i_{k-1}}(n_k)\,p_{i_{k+1}i_{k+1}}^{H,i_1,\dots,i_k}(n_{k+1})\\
&\quad=\sum_{m=0}^{n}\sum_{n_1+\cdots+n_{k-1}+n_{k+2}+\cdots+n_s=n-m} p_{i_1i_1}^{H}(n_1)\cdots p_{i_{k-1}i_{k-1}}^{H,i_1,\dots,i_{k-2}}(n_{k-1})\\
&\qquad\times p_{i_{k+2}i_{k+2}}^{H,i_1,\dots,i_{k+1}}(n_{k+2})\cdots p_{i_si_s}^{H,i_1,\dots,i_{s-1}}(n_s)\\
&\qquad\times\sum_{n_k+n_{k+1}=m} p_{i_{k+1}i_{k+1}}^{H,i_1,\dots,i_{k-1}}(n_k)\,p_{i_ki_k}^{H,i_1,\dots,i_{k-1},i_{k+1}}(n_{k+1})\\
&\quad=\sum_{n_1+\cdots+n_s=n} p_{i_1i_1}^{H}(n_1)\cdots p_{i_{k-1}i_{k-1}}^{H,i_1,\dots,i_{k-2}}(n_{k-1})\,p_{i_{k+1}i_{k+1}}^{H,i_1,\dots,i_{k-1}}(n_k)\\
&\qquad\times p_{i_ki_k}^{H,i_1,\dots,i_{k-1},i_{k+1}}(n_{k+1})\,p_{i_{k+2}i_{k+2}}^{H,i_1,\dots,i_{k+1}}(n_{k+2})\cdots p_{i_si_s}^{H,i_1,\dots,i_{s-1}}(n_s)\\
&\quad=G_n^H(i_1,\dots,i_{k-1},i_{k+1},i_k,i_{k+2},\dots,i_s).
\end{aligned}$$

This completes the proof of this lemma. □

The following lemma will play a key role in the proof of the generalized Haldane equality.

LEMMA 3.10. *Let $c_1, c_2, \dots, c_r$ be a family of cycles passing through a common state i. Let $T = \min\{T^{c_1}, T^{c_2}, \dots, T^{c_r}\}$. Let $c_k = (i, i_2^k, \dots, i_s^k)$. Then for each $n \geq 1$,*

$$P_i\big(T^{c_k} = n, T = T^{c_k}\big) = F_n^i(i_2^k,\dots,i_s^k)\gamma^{c_k}, \tag{28}$$

where $F_n^i(i_2^k,\dots,i_s^k)$ is invariant under any permutation of $i_2^k,\dots,i_s^k$.

PROOF. Note that the event $\{T^{c_k} = n, T = T^{c_k}\}$ is equivalent to saying that X forms c_k at time n and does not form $c_1, c_2, \dots, c_r$ before time n. In order to make this event occur, the Markov chain X must finish the following procedures.

First, X must take n_1 steps to return from i to i without forming $c_1, c_2, \dots, c_r$, and then jump from i to i_2^k. Second, X must take n_2 steps to return from i_2^k to i_2^k without entering i and without forming $c_1, c_2, \dots, c_r$, and then jump from i_2^k to i_3^k. Third, X must take n_3 steps to return from i_3^k to i_3^k without entering i, i_2^k and without forming $c_1, c_2, \dots, c_r$, and then jump from i_3^k to i_4^k, and so on. Finally, X must take n_s steps to return from i_s^k to i_s^k without entering $i, i_1^k, \dots, i_{s-1}^k$ and without forming $c_1, c_2, \dots, c_r$, and then jump from i_s^k to i. Here, the steps $n_1, n_2, \dots, n_s$ must satisfy $(n_1+1)+(n_2+1)+\cdots+(n_s+1)=n$, that is, $n_1+n_2+\cdots+n_s=n-s$.

We make a crucial observation that if X does not enter i, it will not form any one of $c_1, c_2, \dots, c_r$ since all these cycles pass through i. Let $p_{ii}^{c_1,c_2,\dots,c_r}(n_1)$ denote the probability that X takes n_1 steps to return from i to i without forming $c_1, c_2, \dots, c_r$. According to the above discussion, we obtain that

$$
\begin{aligned}
P_i\bigl(T^{c_k}=n, T=T^{c_k}\bigr)
&= \sum_{n_1+n_2+\cdots+n_s=n-s} p_{ii}^{c_1,c_2,\dots,c_r}(n_1)\, p_{i i_2^k}\, p_{i_2^k i_2^k}^{i}(n_2)\, p_{i_2^k i_3^k}\, p_{i_3^k i_3^k}^{i,i_2^k}(n_3) \\
&\quad \times p_{i_3^k i_4^k} \cdots p_{i_s^k i_s^k}^{i,i_1^k,\dots,i_{s-1}^k}(n_s)\, p_{i_s^k i} \\
&= \Biggl[\sum_{n_1=0}^{n-s} p_{ii}^{c_1,c_2,\dots,c_r}(n_1)\, G_{n-n_1-s}^{i}\bigl(i_2^k, \dots, i_s^k\bigr)\Biggr] p_{i i_2^k}\, p_{i_2^k i_3^k} \cdots p_{i_s^k i},
\end{aligned}
$$

where

$$
\begin{aligned}
&G_{n-n_1-s}^{i}\bigl(i_2^k, \dots, i_s^k\bigr) \\
&\qquad = \sum_{n_2+\cdots+n_s=n-n_1-s} p_{i_2^k i_2^k}^{i}(n_2)\, p_{i_3^k i_3^k}^{i,i_2^k}(n_3) \cdots p_{i_s^k i_s^k}^{i,i_1^k,\dots,i_{s-1}^k}(n_s).
\end{aligned}
\tag{29}
$$

By Lemma 3.9, $G_{n-n_1-s}^{i}(i_2^k, \dots, i_s^k)$ is invariant under any permutation of $i_2^k, \dots, i_s^k$. Let

$$
F_n^i\bigl(i_2^k, \dots, i_s^k\bigr) = \sum_{n_1=0}^{n-s} p_{ii}^{c_1,c_2,\dots,c_r}(n_1)\, G_{n-n_1-s}^{i}\bigl(i_2^k, \dots, i_s^k\bigr). \tag{30}
$$

Then $F_n^i(i_2^k, \dots, i_s^k)$ is invariant under any permutation of $i_2^k, \dots, i_s^k$. This completes the proof of this lemma. $\square$

REMARK 3.11. The core idea in the above proof is to decompose the state transitions of each trajectory in the event $\{T^{c_k}=n, T=T^{c_k}\}$ into invalid transitions and valid transitions. During the invalid transitions, X will walk around in circles without contributing to the forming of cycle c_k. During the valid transitions, however, X will jump along cycle c_k. In this way, we can decompose the probability $P_i(T^{c_k} = n, T = T^{c_k})$ into the product of an invalid part $F_n^i(i_2^k, \dots, i_s^k)$ and a valid part γ^{c_k}. The invalid part is invariant under any permutation of $i_2^k, \dots, i_s^k$ and the valid part is independent of time n.

We are now in a position to prove the generalized Haldane equality.

PROOF OF THEOREM 3.3. It is easy to see that (ii) is a direct corollary of (i). Thus, we only need to prove (i). Without loss of generality, we assume that X starts from an arbitrarily fixed state i. Since $c_1, c_2, \dots, c_r$ are similar, they must pass through the same set of states.

We first consider the case when $c_1, c_2, \dots, c_r$ pass through i. Let $c_k = (i, i_2^k, \dots, i_s^k)$ and $c_l = (i, i_2^l, \dots, i_s^l)$. By Lemma 3.10, we have

$$\begin{aligned} P_i\bigl(T^{c_k} = n, T = T^{c_k}\bigr) &= F_n^i\bigl(i_2^k, \dots, i_s^k\bigr)\gamma^{c_k}, \\ P_i\bigl(T^{c_l} = n, T = T^{c_l}\bigr) &= F_n^i\bigl(i_2^l, \dots, i_s^l\bigr)\gamma^{c_l}, \end{aligned} \tag{31}$$

where $F_n^i(i_2^k, \dots, i_s^k)$ is invariant under any permutation of $i_2^k, \dots, i_s^k$. Since c_k and c_l are similar, $i_2^k, \dots, i_s^k$ can be transformed into $i_2^l, \dots, i_s^l$ by a permutation. This shows that

$$\frac{P_i(T^{c_k} = n, T = T^{c_k})}{P_i(T^{c_l} = n, T = T^{c_l})} = \frac{\gamma^{c_k}}{\gamma^{c_l}}. \tag{32}$$

We next consider the case when $c_1, c_2, \dots, c_r$ do not pass through i. Let $c_k = (i_1^k, i_2^k, \dots, i_s^k)$ and $c_l = (i_1^l, i_2^l, \dots, i_s^l)$. By an argument similar to the proof of Lemma 3.10, we can obtain that

$$\begin{aligned} &P_i\bigl(T^{c_k} = n, T = T^{c_k}\bigr) \\ &\quad = \sum_{m=1}^{s} \sum_{n_1+n_2+\cdots+n_s=n-s} p_{i,i_m^k}^{c_1,c_2,\dots,c_r}(n_1) p_{i_m^k i_{m+1}^k} p_{i_{m+1}^k i_{m+1}^k}^{i_m^k}(n_2) \\ &\qquad \times p_{i_{m+1}^k i_{m+2}^k} \cdots p_{i_{m+s-1}^k i_{m+s-1}^k}^{i_m^k,\dots,i_{m+s-2}^k}(n_s) p_{i_{m+s-1}^k i_m^k} \\ &\quad = \gamma^{c_k} \sum_{m=1}^{s} \sum_{n_1=0}^{n-s} p_{i,i_m^k}^{c_1,c_2,\dots,c_r}(n_1) G_{n-n_1-s}^{i_m^k}\bigl(i_{m+1}^k, \dots, i_{m+s-1}^k\bigr), \end{aligned}$$

where $p_{ii_m^k}^{c_1,c_2,\dots,c_r}(n_1)$ denotes the probability that X takes n_1 steps to jump from i to i_m^k without forming $c_1, c_2, \dots, c_r$ and

$$\begin{aligned} &G_{n-n_1-s}^{i_m^k}\bigl(i_{m+1}^k, \dots, i_{m+s-1}^k\bigr) \\ &\quad = \sum_{n_2+\cdots+n_s=n-n_1-s} p_{i_{m+1}^k i_{m+1}^k}^{i_m^k}(n_2) \cdots p_{i_{m+s-1}^k i_{m+s-1}^k}^{i_m^k,\dots,i_{m+s-2}^k}(n_s). \end{aligned} \tag{33}$$

By Lemma 3.9, it is not difficult to see that

$$\sum_{m=1}^{s}\sum_{n_1=0}^{n-s} p_{i,i_m^k}^{c_1,c_2,\dots,c_r}(n_1) G_{n-n_1-s}^{i_m^k}\big(i_{m+1}^k,\dots,i_{m+s-1}^k\big)$$

is invariant under any permutation of $i_1^k, i_2^k, \dots, i_s^k$. Since c_k and c_l are similar, $i_1^k, i_2^k, \dots, i_s^k$ can be transformed into $i_1^l, i_2^l, \dots, i_s^l$ by a permutation. This suggests that

$$\frac{P_i(T^{c_k}=n, T=T^{c_k})}{P_i(T^{c_l}=n, T=T^{c_l})} = \frac{\gamma^{c_k}}{\gamma^{c_l}}, \tag{34}$$

which gives the first equality in (16). Since the above equation holds for any $n \geq 1$ and the right-hand side of the above equation does not depend on n, the second equality in (16) also holds. □

The next corollary gives another aspect of the generalized Haldane equality.

COROLLARY 3.12. *Let $c_1, c_2, \dots, c_r$ be a family of similar cycles. Let $T = \min\{T^{c_1}, T^{c_2}, \dots, T^{c_r}\}$. Then for each $n \geq 0$ and $1 \leq k \leq r$,*

$$P\big(T=n, T=T^{c_k}\big) = P(T=n)P\big(T=T^{c_k}\big). \tag{35}$$

PROOF. By Theorem 3.3, it is easy to see that for each $n \geq 0$ and $1 \leq k \leq r$,

$$P\big(T=n|T=T^{c_k}\big) = P(T=n). \tag{36}$$

Thus, we obtain that

$$\begin{aligned} P\big(T=n, T=T^{c_k}\big) &= P\big(T=n|T=T^{c_k}\big)P\big(T=T^{c_k}\big) \\ &= P(T=n)P\big(T=T^{c_k}\big), \end{aligned} \tag{37}$$

which gives the desired result. □

REMARK 3.13. The notation is the same as in Corollary 3.12. Let ξ be a random variable defined by

$$\xi = \begin{cases} c_1, & \text{if the trajectory of } X \text{ forms cycle } c_1 \text{ at time } T, \\ c_2, & \text{if the trajectory of } X \text{ forms cycle } c_2 \text{ at time } T, \\ \dots, & \\ c_r, & \text{if the trajectory of } X \text{ forms cycle } c_r \text{ at time } T. \end{cases} \tag{38}$$

It is easy to see that $\xi = c_k$ if and only if $T = T^{c_k}$ for each $1 \leq k \leq r$. Thus, Corollary 3.12 is equivalent to saying that T and ξ are independent. This suggests that the forming time of multiple similar cycles is independent of which one of these cycles is formed. This is another important aspect of the generalized Haldane equality.

In applications, we are more concerned about the symmetry of a cycle and its reversed cycle. Thus, we introduce the following definition.

DEFINITION 3.14. Let $c = (i_1, i_2, \ldots, i_s)$ be a cycle. Then the reversed cycle of cycle c is defined as $c- = (i_1, i_s, \ldots, i_2)$.

It is easy to see that a cycle c and its reversed cycle $c-$ must be similar. Thus, we obtain the following corollary.

COROLLARY 3.15. *Let $c = (i_1, i_2, \ldots, i_s)$ be a cycle. Then*:

(i) *for each $n \geq 0$,*

$$\frac{P(T^c = n, T^c < T^{c-})}{P(T^{c-} = n, T^{c-} < T^c)} = \frac{P(T^c < T^{c-})}{P(T^{c-} < T^c)} = \frac{p_{i_1 i_2} p_{i_2 i_3} \cdots p_{i_s i_1}}{p_{i_1 i_s} p_{i_s i_{s-1}} \cdots p_{i_2 i_1}}; \tag{39}$$

(ii) *for each $n \geq 0$,*

$$P\big(T^c = n | T^c < T^{c-}\big) = P\big(T^{c-} = n | T^{c-} < T^c\big); \tag{40}$$

(iii) *for each $n \geq 0$,*

$$P\big(T^c \wedge T^{c-} = n, T^c < T^{c-}\big) = P\big(T^c \wedge T^{c-} = n\big) P\big(T^c < T^{c-}\big). \tag{41}$$

PROOF. This corollary follows directly from Theorem 3.3 and Corollary 3.12. □

REMARK 3.16. The above corollary generalizes the so-called generalized Haldane equality [see (4) in Section 1] found by biophysicists in cyclic Markov chains [16, 18, 33].

4. Generalizations of the generalized Haldane equality. We have seen that the most important step in the proof of the generalized Haldane equality is Lemma 3.10, in which we decompose the probability $P_i(T^{c_k} = n, T = T^{c_k})$ into an invalid part and a valid part. However, the conditions of Lemma 3.10 are much weaker than those of Theorem 3.3. This suggests that the generalized Haldane equality can be further generalized, as stated in the following theorem.

THEOREM 4.1. *Let $c_1, c_2, \ldots, c_r$ be a family of cycles passing through a common state i. Let $T = \min\{T^{c_1}, T^{c_2}, \ldots, T^{c_r}\}$. Assume that c_k and c_l are similar for some two indices $1 \leq k, l \leq r$. Then*:

(i) *for each $n \geq 1$,*

$$\frac{P_i(T^{c_k} = n, T = T^{c_k})}{P_i(T^{c_l} = n, T = T^{c_l})} = \frac{P_i(T = T^{c_k})}{P_i(T = T^{c_l})} = \frac{\gamma^{c_k}}{\gamma^{c_l}}; \tag{42}$$

(ii) *for each* $n \geq 1$,

$$P_i\big(T^{c_k}=n|T=T^{c_k}\big)=P_i\big(T^{c_l}=n|T=T^{c_l}\big). \tag{43}$$

PROOF. It is easy to see that (ii) is a direct corollary of (i). Thus, we only need to prove (i). Let $c_k=(i,i_2^k,\ldots,i_s^k)$ and $c_l=(i,i_2^l,\ldots,i_s^l)$. By Lemma 3.10, we have

$$\begin{aligned} P_i\big(T^{c_k}=n,T=T^{c_k}\big)&=F_n^i(i_2^k,\ldots,i_s^k)\gamma^{c_k},\\ P_i\big(T^{c_l}=n,T=T^{c_l}\big)&=F_n^i(i_2^l,\ldots,i_s^l)\gamma^{c_l}, \end{aligned} \tag{44}$$

where $F_n^i(i_2^k,\ldots,i_s^k)$ is invariant under any permutation of $i_2^k,\ldots,i_s^k$. Since c_k and c_l are similar, $i_2^k,\ldots,i_s^k$ can be transformed into $i_2^l,\ldots,i_s^l$ by a permutation. This shows that

$$\frac{P_i(T^{c_k}=n,T=T^{c_k})}{P_i(T^{c_l}=n,T=T^{c_l})}=\frac{\gamma^{c_k}}{\gamma^{c_l}}. \tag{45}$$

This completes the proof of this theorem. □

REMARK 4.2. There are two crucial differences between Theorems 3.3 and 4.1. The first difference is that we require $c_1,c_2,\ldots,c_r$ to be similar in Theorem 3.3, while we only require $c_1,c_2,\ldots,c_r$ to pass through a common state in Theorem 4.1. The second difference is that Theorem 3.3 holds for Markov chains starting from any initial distributions, while Theorem 4.1 only holds for Markov chains starting from the common state.

Since a cycle c and its reversed cycle $c-$ must be similar, we obtain the following corollary.

COROLLARY 4.3. *Let* $c_1,c_2,\ldots,c_r$ *be a family of cycles passing through a common state* i. *Let* $T=\min\{T^{c_1},T^{c_1-},\ldots,T^{c_r},T^{c_r-}\}$. *Then*:

(i) *for each* $n\geq 1$ *and* $1\leq k\leq r$,

$$\frac{P_i(T^{c_k}=n,T=T^{c_k})}{P_i(T^{c_k-}=n,T=T^{c_k-})}=\frac{P_i(T=T^{c_k})}{P_i(T=T^{c_k-})}=\frac{\gamma^{c_k}}{\gamma^{c_k-}}; \tag{46}$$

(ii) *for each* $n\geq 1$ *and* $1\leq k\leq r$,

$$P_i\big(T^{c_k}=n|T=T^{c_k}\big)=P_i\big(T^{c_k-}=n|T=T^{c_k-}\big). \tag{47}$$

PROOF. This corollary follows directly from Theorem 4.1. □

REMARK 4.4. We have seen that the generalized Haldane equality has many variations which are closely related. These results, which include Theorems 3.3 and 4.1, Corollaries 3.12, 3.15 and 4.3, will be collectively referred to as the generalized Haldane equalities in the following discussion.

5. Generalized Haldane equalities for continuous-time Markov chains. In this section, we shall prove the generalized Haldane equalities for continuous-time Markov chains. Let $X = (X_t)_{t\geq 0}$ be an irreducible and recurrent continuous-time Markov chain with denumerable state space S and transition rate matrix $Q = (q_{ij})$.

Before we state the generalized Haldane equality, we give the definition of the strengths of cycles for continuous-time Markov chains.

DEFINITION 5.1. Let $c = (i_1, i_2, \dots, i_s)$ be a cycle. Then the strength of cycle c is defined as

$$\gamma^c = q_{i_1 i_2} q_{i_2 i_3} \cdots q_{i_s i_1}. \tag{48}$$

The generalized Haldane equality, which characterizes the symmetry of the forming times of a family of similar cycles, is stated in the following theorem.

THEOREM 5.2. *Let $c_1, c_2, \dots, c_r$ be a family of similar cycles. Let $T = \min\{T^{c_1}, T^{c_2}, \dots, T^{c_r}\}$. Then*:

(i) *for each $t > 0$ and any $1 \leq k, l \leq r$,*

$$\frac{P(T^{c_k} \leq t, T = T^{c_k})}{P(T^{c_l} \leq t, T = T^{c_l})} = \frac{P(T = T^{c_k})}{P(T = T^{c_l})} = \frac{\gamma^{c_k}}{\gamma^{c_l}}; \tag{49}$$

(ii) *for each $t > 0$,*

$$\begin{aligned} P\big(T^{c_1} \leq t | T = T^{c_1}\big) &= P\big(T^{c_2} \leq t | T = T^{c_2}\big) \\ &= \cdots = P\big(T^{c_r} \leq t | T = T^{c_r}\big). \end{aligned} \tag{50}$$

PROOF. It is easy to see that (ii) is a direct corollary of (i). Thus, we only need to prove (i). Let $t > 0$ be an arbitrarily fixed time. For each $m \geq 1$, let

$$Y_n^m = X_{nt/m}. \tag{51}$$

Then $Y^m = (Y_n^m)_{n\geq 0}$ is an irreducible and recurrent discrete-time Markov chain with transition probability matrix $P_m = (p_{ij}(t/m))$, where $p_{ij}(t/m) = P_i(X_{t/m} = j)$. Let $T^{m,c}$ be the forming time of cycle c by Y^m. Let $T^m = \min\{T^{m,c_1}, T^{m,c_2}, \dots, T^{m,c_r}\}$.

Since X is irreducible and recurrent, it must be nonexplosive, which implies that X can only complete a finite number of jumps by time t. Thus, for any $\omega \in \Omega$, when m is sufficiently large, t/m is less than any of the waiting times of $X(\omega)$ by time t. This means that the occurrence of the event $\{T^{c_k} \leq t, T = T^{c_k}\}$ implies the occurrence of the event $\{T^{m,c_k} \leq m, T^m = T^{m,c_k}\}$ when m is sufficiently large. Thus, we obtain that

$$\{T^{c_k} \leq t, T = T^{c_k}\} \subset \bigcup_{N=1}^{\infty} \bigcap_{m=N}^{\infty} \{T^{m,c_k} \leq m, T^m = T^{m,c_k}\}. \tag{52}$$

Similarly, it is easy to see that the occurrence of the event $\{T^{c_k} > t\}$ implies the occurrence of the event $\{T^{m,c_k} > m\}$ when m is sufficiently large, and the occurrence of the event $\{T < T^{c_k} \leq t\}$ implies the occurrence of the event $\{T^m < T^{m,c_k} \leq m\}$ when m is sufficiently large. Thus, we obtain that

$$
\begin{aligned}
&\{T^{c_k} \leq t, T = T^{c_k}\}^c \\
&\quad = \{T^{c_k} > t\} \cup \{T < T^{c_k} \leq t\} \\
&\quad \subset \left(\bigcup_{N=1}^{\infty} \bigcap_{m=N}^{\infty} \{T^{m,c_k} > m\}\right) \cup \left(\bigcup_{N=1}^{\infty} \bigcap_{m=N}^{\infty} \{T^m < T^{m,c_k} \leq m\}\right) \qquad (53) \\
&\quad \subset \bigcup_{N=1}^{\infty} \bigcap_{m=N}^{\infty} \{T^{m,c_k} > m\} \cup \{T^m < T^{m,c_k} \leq m\} \\
&\quad = \bigcup_{N=1}^{\infty} \bigcap_{m=N}^{\infty} \{T^{m,c_k} \leq m, T^m = T^{m,c_k}\}^c.
\end{aligned}
$$

This shows that

$$
\bigcap_{N=1}^{\infty} \bigcup_{m=N}^{\infty} \{T^{m,c_k} \leq m, T^m = T^{m,c_k}\} \subset \{T^{c_k} \leq t, T = T^{c_k}\}. \qquad (54)
$$

By (52) and (54), we have

$$
\{T^{c_k} \leq t, T = T^{c_k}\} = \lim_{m\to\infty} \{T^{m,c_k} \leq m, T^m = T^{m,c_k}\}. \qquad (55)
$$

By the dominated convergence theorem, we obtain that

$$
P(T^{c_k} \leq t, T = T^{c_k}) = \lim_{m\to\infty} P(T^{m,c_k} \leq m, T^m = T^{m,c_k}). \qquad (56)
$$

Let $c_k = (i_1^k, i_2^k, \dots, i_s^k)$ and $c_l = (i_1^l, i_2^l, \dots, i_s^l)$. By Theorem 3.3, we have

$$
\begin{aligned}
\frac{P(T^{c_k} \leq t, T = T^{c_k})}{P(T^{c_l} \leq t, T = T^{c_l})} &= \lim_{m\to\infty} \frac{P(T^{m,c_k} \leq m, T^m = T^{m,c_k})}{P(T^{m,c_l} \leq m, T^m = T^{m,c_l})} \\
&= \lim_{m\to\infty} \frac{p_{i_1^k i_2^k}(t/m) p_{i_2^k i_3^k}(t/m) \cdots p_{i_s^k i_1^k}(t/m)}{p_{i_1^l i_2^l}(t/m) p_{i_2^l i_3^l}(t/m) \cdots p_{i_s^l i_1^l}(t/m)} \qquad (57) \\
&= \frac{q_{i_1^k i_2^k} q_{i_2^k i_3^k} \cdots q_{i_s^k i_1^k}}{q_{i_1^l i_2^l} q_{i_2^l i_3^l} \cdots q_{i_s^l i_1^l}} = \frac{\gamma^{c_k}}{\gamma^{c_l}}.
\end{aligned}
$$

This completes the proof of this theorem. $\square$

We can also obtain the following results parallel to those for discrete-time Markov chains.

COROLLARY 5.3. *Let $c_1, c_2, \dots, c_r$ be a family of similar cycles. Let $T = \min\{T^{c_1}, T^{c_2}, \dots, T^{c_r}\}$. Then for each $t > 0$ and $1 \le k \le r$,*

$$P(T \le t, T = T^{c_k}) = P(T \le t)P(T = T^{c_k}). \tag{58}$$

PROOF. The proof of this corollary follows the same lines as that of Corollary 3.12. □

COROLLARY 5.4. *Let $c = (i_1, i_2, \dots, i_s)$ be a cycle. Then*:

(i) *for each $t > 0$,*

$$\frac{P(T^c \le t, T^c < T^{c-})}{P(T^{c-} \le t, T^{c-} < T^c)} = \frac{P(T^c < T^{c-})}{P(T^{c-} < T^c)} = \frac{q_{i_1 i_2} q_{i_2 i_3} \cdots q_{i_s i_1}}{q_{i_1 i_s} q_{i_s i_{s-1}} \cdots q_{i_2 i_1}}; \tag{59}$$

(ii) *for each $t > 0$,*

$$P(T^c \le t | T^c < T^{c-}) = P(T^{c-} \le t | T^{c-} < T^c); \tag{60}$$

(iii) *for each $t > 0$,*

$$P(T^c \wedge T^{c-} \le t, T^c < T^{c-}) = P(T^c \wedge T^{c-} \le t)P(T^c < T^{c-}). \tag{61}$$

PROOF. This corollary follows directly from Theorem 5.2 and Corollary 5.3. □

THEOREM 5.5. *Let $c_1, c_2, \dots, c_r$ be a family of cycles passing through a common state i. Let $T = \min\{T^{c_1}, T^{c_2}, \dots, T^{c_r}\}$. Assume that c_k and c_l are similar for some two indices $1 \le k, l \le r$. Then*:

(i) *for each $t > 0$,*

$$\frac{P_i(T^{c_k} \le t, T = T^{c_k})}{P_i(T^{c_l} \le t, T = T^{c_l})} = \frac{P_i(T = T^{c_k})}{P_i(T = T^{c_l})} = \frac{\gamma^{c_k}}{\gamma^{c_l}}; \tag{62}$$

(ii) *for each $t > 0$,*

$$P_i(T^{c_k} \le t | T = T^{c_k}) = P_i(T^{c_l} \le t | T = T^{c_l}). \tag{63}$$

PROOF. It is easy to see that (ii) is a direct corollary of (i). Thus, we only need to prove (i). Let $t > 0$ be an arbitrarily fixed time. For each $m \ge 1$, let

$$Y_n^m = X_{nt/m}. \tag{64}$$

Then $Y^m = (Y_n^m)_{n \ge 0}$ is an irreducible and recurrent discrete-time Markov chain with transition probability matrix $P_m = (p_{ij}(t/m))$, where $p_{ij}(t/m) = P_i(X_{t/m} = j)$. Let $T^{m,c}$ be the forming time of cycle c by Y^m. Let $T^m = \min\{T^{m,c_1}, T^{m,c_2}, \dots, T^{m,c_r}\}$.

By an argument similar to the proof of Theorem 5.2, we can obtain that

$$\{T^{c_k} \le t, T = T^{c_k}\} = \lim_{m\to\infty} \{T^{m,c_k} \le m, T^m = T^{m,c_k}\}. \tag{65}$$

By the dominated convergence theorem, we obtain that

$$P_i(T^{c_k} \le t, T = T^{c_k}) = \lim_{m\to\infty} P_i(T^{m,c_k} \le m, T^m = T^{m,c_k}). \tag{66}$$

Let $c_k = (i_1^k, i_2^k, \dots, i_s^k)$ and $c_l = (i_1^l, i_2^l, \dots, i_s^l)$. By Theorem 3.3, we have

$$\begin{aligned} \frac{P_i(T^{c_k} \le t, T = T^{c_k})}{P_i(T^{c_l} \le t, T = T^{c_l})} &= \lim_{m\to\infty} \frac{P_i(T^{m,c_k} \le m, T^m = T^{m,c_k})}{P_i(T^{m,c_l} \le m, T^m = T^{m,c_l})} \\ &= \lim_{m\to\infty} \frac{p_{i_1^k i_2^k}(t/m) p_{i_2^k i_3^k}(t/m) \cdots p_{i_s^k i_1^k}(t/m)}{p_{i_1^l i_2^l}(t/m) p_{i_2^l i_3^l}(t/m) \cdots p_{i_s^l i_1^l}(t/m)} \\ &= \frac{q_{i_1^k i_2^k} q_{i_2^k i_3^k} \cdots q_{i_s^k i_1^k}}{q_{i_1^l i_2^l} q_{i_2^l i_3^l} \cdots q_{i_s^l i_1^l}} = \frac{\gamma^{c_k}}{\gamma^{c_l}}. \end{aligned} \tag{67}$$

This completes the proof of this theorem. $\square$

COROLLARY 5.6. *Let $c_1, c_2, \dots, c_r$ be a family of cycles passing through a common state i. Let $T = \min\{T^{c_1}, T^{c_1-}, \dots, T^{c_r}, T^{c_r-}\}$. Then*:

(i) *for each $t > 0$ and $1 \le k \le r$,*

$$\frac{P_i(T^{c_k} \le t, T = T^{c_k})}{P_i(T^{c_k-} \le t, T = T^{c_k-})} = \frac{P_i(T = T^{c_k})}{P_i(T = T^{c_k-})} = \frac{\gamma^{c_k}}{\gamma^{c_k-}}; \tag{68}$$

(ii) *for each $t > 0$ and $1 \le k \le r$,*

$$P_i(T^{c_k} \le t | T = T^{c_k}) = P_i(T^{c_k-} \le t | T = T^{c_k-}). \tag{69}$$

PROOF. This corollary follows directly from Theorem 5.5. $\square$

6. Large deviations and fluctuations of sample circulations. The generalized Haldane equalities established in the above sections have wide applications. One of the most important applications is to study the circulation fluctuations for Markov chains. In this section, we shall prove that the sample circulations along a family of cycles passing through a common state satisfy a large deviation principle with a good rate function. Particularly, we shall use the generalized Haldane equalities to prove that the rate function has a highly nonobvious symmetry, which is closely related to the Gallavotti–Cohen-type fluctuation theorem in nonequilibrium statistical physics.

6.1. *Preliminaries.* In order to study the large deviations of the sample circulations, we need some results about the large deviations of Markov renewal processes. To avoid misunderstanding, we recall the following two definitions.

DEFINITION 6.1. Let $(\mu_t)_{t>0}$ be a family of probability measures on a Polish space E. Then we say that $(\mu_t)_{t>0}$ satisfies a large deviation principle with rate t and good rate function $I : E \to [0, \infty]$ if:

(i) for each $\alpha \geq 0$, the level set $\{x \in E : I(x) \leq \alpha\}$ is compact in E;
(ii) for each closed subset F of E,

$$\limsup_{t\to\infty} \frac{1}{t} \log \mu_t(F) \leq -\inf_{x\in F} I(x); \tag{70}$$

(iii) for each open subset U of E,

$$\liminf_{t\to\infty} \frac{1}{t} \log \mu_t(U) \geq -\inf_{x\in U} I(x). \tag{71}$$

DEFINITION 6.2. Let $\xi = (\xi_n)_{n\geq 0}$ be an irreducible discrete-time Markov chain with finite state space E. Assume that each $x \in E$ is associated with a Borel probability measure ϕ_x on $(0, \infty)$. Let $(\tau_n)_{n\geq 1}$ be a sequence of positive and finite random variables such that conditioned on $(\xi_n)_{n\geq 0}$, the random variables $(\tau_n)_{n\geq 1}$ are independent and have the distribution

$$P\big(\tau_n \in \cdot | (\xi_n)_{n\geq 0}\big) = \phi_{\xi_{n-1}}(\cdot). \tag{72}$$

Then $(\xi_n, \tau_{n+1})_{n\geq 0}$ is called a Markov renewal process.

The following lemma, which is due to Mariani and Zambotti, shows that the empirical flow of Markov renewal processes satisfies a large deviation principle with a good rate function.

LEMMA 6.3. *Let $Z = (\xi_n, \tau_{n+1})_{n\geq 0}$ be a Markov renewal process. Let $T_n = \sum_{k=1}^n \tau_k$ be the nth jump time of Z. Let $N_t = \inf\{n \geq 0 : T_{n+1} > t\}$ be the number of jumps of Z by time t. Let $Q_t \in C(E \times E, [0, \infty))$ be the empirical flow of Z by time t defined as*

$$Q_t(x, y) = \frac{1}{t} \sum_{n=0}^{N_t} 1_{\{\xi_n = x, \xi_{n+1} = y\}}. \tag{73}$$

Then the law of Q_t satisfies a large deviation principle with rate t and good rate function $I : C(E \times E, [0, \infty)) \to [0, \infty]$. Moreover, the rate function I is convex.

PROOF. The proof of this theorem can be found in [29], Theorem 1.2. $\square$

6.2. *Large deviations of sample circulations.* In this paper, we only consider the large deviations of the sample circulations for continuous-time Markov chains. Using similar but simpler techniques, we can obtain parallel results for discrete-time Markov chains.

Let $X = (X_t)_{t\geq 0}$ be an irreducible and recurrent continuous-time Markov chain with denumerable state space S and transition rate matrix $Q = (q_{ij})$. We next give the definition of the sample circulations for Markov chains.

DEFINITION 6.4. Let T_n^c be the nth forming time of cycle c. Let $N_t^c = \inf\{n \geq 0 : T_{n+1}^c > t\}$ be the number of times that cycle c is formed by X up to time t. Then the sample circulation J_t^c along cycle c by time t is defined as

$$J_t^c = \frac{1}{t} N_t^c \tag{74}$$

and the sample net circulation K_t^c along cycle c by time t is defined as $K_t^c = J_t^c - J_t^{c-}$.

We next recall the definition of the circulations for Markov chains.

DEFINITION 6.5. The circulation J^c along cycle c is defined as

$$J^c = \lim_{t\to\infty} J_t^c, \qquad \text{a.s.} \tag{75}$$

and the net circulation K^c along cycle c is defined as $K^c = J^c - J^{c-}$.

Intuitively, J^c represents the number of times that cycle c is formed by X per unit time and K^c represents the net number of times that cycle c formed by X per unit time.

REMARK 6.6. It can be proved that the almost sure limit in (75) exists whenever X is irreducible and recurrent and the limit is a nonnegative constant independent of the initial distribution of X. When X is positive recurrent, the proof of the above fact is due to the Qian's [21, 35]. When X is null recurrent, it is easy to see that $J^c = K^c = 0$ for each cycle c.

We have defined the forming time of a single cycle in Definition 2.11. We shall now define the forming time of multiple cycles.

DEFINITION 6.7. Let $c_1, c_2, \ldots, c_r$ be a family of cycles and let $T_n^{c_1}, T_n^{c_2}, \ldots, T_n^{c_r}$ be their nth forming times. Then the nth forming time T_n of cycles $c_1, c_2, \ldots, c_r$ is defined as the nth order statistics of the set $\{T_m^{c_1}, T_m^{c_2}, \ldots, T_m^{c_r} : m \geq 1\}$.

If we only focus on the forming of cycles by a Markov chain, instead of the specific state transitions, then we can obtain a Markov renewal process, as shown in the following lemma.

LEMMA 6.8. *Let $c_1, c_2, \dots, c_r$ be a family of cycles passing through a common state i and assume that $\gamma^{c_k} > 0$ for some $1 \le k \le r$. Let T_n be the nth forming time of cycles $c_1, c_2, \dots, c_r$. Let $\tau_n = T_n - T_{n-1}$. Let ξ_n be a random variable defined as*

$$\xi_n = \begin{cases} c_1, & \text{if the trajectory of } X \text{ forms cycle } c_1 \text{ at time } T_n, \\ c_2, & \text{if the trajectory of } X \text{ forms cycle } c_2 \text{ at time } T_n, \\ \dots, & \\ c_r, & \text{if the trajectory of } X \text{ forms cycle } c_r \text{ at time } T_n. \end{cases} \tag{76}$$

Then under P_i, $(\xi_n, \tau_n)_{n\ge1}$ is a Markov renewal process.

PROOF. Since $X_0 = i$ and $c_1, c_2, \dots, c_r$ pass through i, it is easy to see that $X_{T_n} = i$ for each $n \ge 1$. By the strong Markov property, the random sequence $(\xi_n, \tau_n)_{n\ge1}$ is an i.i.d. sequence. This shows that $(\xi_n)_{n\ge1}$ is a Markov chain with state space $E = \{c_1, c_2, \dots, c_r\}$. Note that we have assumed that $\gamma^{c_k} > 0$ for some k. By Proposition 2.12, we see that $T_n \le T_n^{c_k} < \infty$, a.s. for each $n \ge 1$. This shows that $(\tau_n)_{n\ge1}$ is a sequence of positive and finite random variables.

Since $(\xi_n, \tau_n)_{n\ge1}$ is an i.i.d. sequence, for any bounded measurable function $f_1, \dots, f_n$ on $(0, \infty)$, it is easy to see that

$$\begin{aligned} &E_i\big(f_1(\tau_1)\cdots f_n(\tau_n)|(\xi_n)_{n\ge1}\big) \\ &\quad = E_i\big(f_1(\tau_1)|(\xi_n)_{n\ge1}\big)\cdots E_i\big(f_n(\tau_n)|(\xi_n)_{n\ge1}\big). \end{aligned} \tag{77}$$

Moreover, for any Borel set A in $(0, \infty)$,

$$P_i\big(\tau_n \in A|(\xi_n)_{n\ge1}\big) = P_i(\tau_n \in A|\xi_n) = P_i(\tau_1 \in A|\xi_1 = x)|_{x=\xi_n} = \phi_{\xi_n}(A), \tag{78}$$

where $\phi_x(A) = P_i(\tau_1 \in A|\xi_1 = x)$. The above two equations show that each $x \in E$ is associated with a Borel probability measure ϕ_x on $(0, \infty)$ and conditioned on $(\xi_n)_{n\ge1}$, the random variables $(\tau_n)_{n\ge1}$ are independent and have the distribution $P_i(\tau_n \in \cdot|(\xi_n)_{n\ge1}) = \phi_{\xi_n}(\cdot)$. This shows that $(\xi_n, \tau_n)_{n\ge1}$ is a Markov renewal process. □

The large deviation principle of the sample circulations is stated in the following theorem.

THEOREM 6.9. *Let $c_1, c_2, \dots, c_r$ be a family of cycles passing through a common state i. Then under P_i, the law of $(J_t^{c_1}, J_t^{c_2}, \dots, J_t^{c_r})$ satisfies a large deviation principle with rate t and good rate function $I^{c_1,c_2,\dots,c_r} : \mathbb{R}^r \to [0, \infty]$.*

PROOF. We first consider the case when $\gamma^{c_k} = 0$ for each k. By Proposition 2.12, we see that $T^{c_k} = \infty$, a.s. for each k. This shows that $J_t^{c_k} = 0$, a.s. for each k. In this case, the theorem holds trivially.

We next consider the case when $\gamma^{c_k} > 0$ for some k. By Lemma 6.8, we see that $(\xi_n, \tau_n)_{n\geq 1}$ is a Markov renewal process with state space $E = \{c_1, c_2, \dots, c_r\}$. Let $N_t = \inf\{n \geq 0 : T_{n+1} > t\}$ be the number of jumps of the Markov renewal process by time t. Let $Q_t \in C(E \times E, [0, \infty))$ be the empirical flow of the Markov renewal process by time t defined as

$$Q_t(x, y) = \frac{1}{t} \sum_{n=1}^{N_t} 1_{\{\xi_n = x, \xi_{n+1} = y\}}. \tag{79}$$

Note that for each k,

$$J_t^{c_k} = \frac{1}{t} N_t^{c_k} = \frac{1}{t} \sum_{n=1}^{N_t} 1_{\{\xi_n = c_k\}} = \sum_{y \in E} Q_t(c_k, y). \tag{80}$$

We define a continuous map $F : C(E \times E, [0, \infty)) \to \mathbb{R}^r$ as

$$F(Q) = \left(\sum_{y \in E} Q(c_1, y), \dots, \sum_{y \in E} Q(c_r, y) \right). \tag{81}$$

Thus, we have

$$\left(J_t^{c_1}, \dots, J_t^{c_r}\right) = F(Q_t). \tag{82}$$

By Lemma 6.3, the law of Q_t satisfies a large deviation principle with rate t and good rate function $I : C(E \times E, [0, \infty)) \to [0, \infty]$. Using the contraction principle, we see that the law of $(J_t^{c_1}, \dots, J_t^{c_r})$ satisfies a large deviation principle with rate t and good rate function $I^{c_1, \dots, c_r} : \mathbb{R}^r \to [0, \infty]$ which can be represented as

$$I^{c_1, \dots, c_r}(x) = \inf_{Q \in F^{-1}(x)} I(Q). \tag{83}$$

This completes the proof of this theorem. $\square$

6.3. *Circulation fluctuations for Markov chains*. We have proved that the sample circulations along a family of cycles $c_1, c_2, \dots, c_r$ passing through a common state satisfy a large deviation principle with rate t and good rate function $I^{c_1, c_2, \dots, c_r}$. In general, it is very difficult to obtain an explicit expression of the rate function $I^{c_1, c_2, \dots, c_r}$. However, we can use the generalized Haldane equalities to prove that the rate function $I^{c_1, c_2, \dots, c_r}$ has a highly nonobvious symmetry, whose specific form is given in the following theorem.

THEOREM 6.10. *The notation is the same as in Theorem 6.9. Assume that c_k and c_l are similar for some two indices $1 \le k, l \le r$. Then the rate function $I^{c_1,c_2,\dots,c_r}$ has the following symmetry: for any $x_1, x_2, \dots, x_r \in \mathbb{R}$,*

$$\begin{aligned} &I^{c_1,c_2,\dots,c_r}(x_1,\dots,x_k,\dots,x_l,\dots,x_r) \\ &\quad = I^{c_1,c_2,\dots,c_r}(x_1,\dots,x_l,\dots,x_k,\dots,x_r) - \left(\log \frac{\gamma^{c_k}}{\gamma^{c_l}}\right)(x_k - x_l). \end{aligned} \tag{84}$$

The above theorem shows that if c_k and c_l are similar, then the rate function $I^{c_1,c_2,\dots,c_r}$ satisfies a symmetric relation under the exchange of the arguments x_k and x_l. In order to prove the above theorem, we need several lemmas.

LEMMA 6.11. *The rate function $I^{c_1,c_2,\dots,c_r}$ is convex.*

PROOF. Note that the map F defined in (81) is linear. This fact, together with (83), shows that for each $0 < \lambda < 1$ and any $x, y \in \mathbb{R}^r$,

$$\begin{aligned} I^{c_1,\dots,c_r}\big(\lambda x + (1-\lambda)y\big) &= \inf_{Q \in F^{-1}(\lambda x + (1-\lambda)y)} I(Q) \\ &\le \inf_{Q \in \lambda F^{-1}(x) + (1-\lambda)F^{-1}(y)} I(Q) \\ &= \inf_{Q \in F^{-1}(x),\, R \in F^{-1}(y)} I\big(\lambda Q + (1-\lambda)R\big). \end{aligned} \tag{85}$$

By Lemma 6.3, the rate function I is convex. Thus, we obtain that

$$\begin{aligned} I^{c_1,\dots,c_r}\big(\lambda x + (1-\lambda)y\big) &\le \inf_{Q \in F^{-1}(x),\, R \in F^{-1}(y)} \lambda I(Q) + (1-\lambda) I(R) \\ &= \lambda \inf_{Q \in F^{-1}(x)} I(Q) + (1-\lambda) \inf_{R \in F^{-1}(y)} I(R) \\ &= \lambda I^{c_1,\dots,c_r}(x) + (1-\lambda) I^{c_1,\dots,c_r}(y). \end{aligned} \tag{86}$$

This completes the proof of this lemma. □

The following lemma follows directly from the generalized Haldane equalities.

LEMMA 6.12. *Let $c_1, c_2, \dots, c_r$ be a family of cycles passing through a common state i. Assume that c_k and c_l are similar for some two indices $1 \le k, l \le r$. Let $T = \min\{T^{c_1}, T^{c_2}, \dots, T^{c_r}\}$. Then for each $t > 0$,*

$$P_i\big(T \le t, T = T^{c_k}\big) = P_i\big(T \le t, T = T^{c_k} \wedge T^{c_l}\big) P_i\big(T^{c_k} < T^{c_l}\big). \tag{87}$$

PROOF. By Theorem 5.5(i), we have

$$\frac{P_i(T \le t, T = T^{c_k})}{P_i(T \le t, T = T^{c_l})} = \frac{P_i(T \le t, T = T^{c_k} \wedge T^{c_l}, T^{c_k} < T^{c_l})}{P_i(T \le t, T = T^{c_k} \wedge T^{c_l}, T^{c_l} < T^{c_k})} = \frac{\gamma^{c_k}}{\gamma^{c_l}}. \tag{88}$$

Using Theorem 5.5(i) again, we have

$$\frac{P_i(T^{c_k} \wedge T^{c_l} = T^{c_k})}{P_i(T^{c_k} \wedge T^{c_l} = T^{c_l})} = \frac{P_i(T^{c_k} < T^{c_l})}{P_i(T^{c_l} < T^{c_k})} = \frac{\gamma^{c_k}}{\gamma^{c_l}}. \tag{89}$$

Combining the above two equations, we obtain that

$$\begin{aligned} &P_i\bigl(T \le t, T = T^{c_k} \wedge T^{c_l} | T^{c_k} < T^{c_l}\bigr) \\ &\qquad = P_i\bigl(T \le t, T = T^{c_k} \wedge T^{c_l} | T^{c_l} < T^{c_k}\bigr). \end{aligned} \tag{90}$$

This implies that

$$P_i\bigl(T \le t, T = T^{c_k} \wedge T^{c_l} | T^{c_k} < T^{c_l}\bigr) = P_i\bigl(T \le t, T = T^{c_k} \wedge T^{c_l}\bigr). \tag{91}$$

Thus, we obtain that

$$\begin{aligned} P_i\bigl(T \le t, T = T^{c_k}\bigr) &= P_i\bigl(T \le t, T = T^{c_k} \wedge T^{c_l}, T^{c_k} < T^{c_l}\bigr) \\ &= P_i\bigl(T \le t, T = T^{c_k} \wedge T^{c_l} | T^{c_k} < T^{c_l}\bigr) P_i\bigl(T^{c_k} < T^{c_l}\bigr) \\ &= P_i\bigl(T \le t, T = T^{c_k} \wedge T^{c_l}\bigr) P_i\bigl(T^{c_k} < T^{c_l}\bigr), \end{aligned} \tag{92}$$

which gives the desired result. $\square$

The following lemma, whose proof strongly depends on the generalized Haldane equalities, characterizes the symmetry of the joint distribution of the sample circulations.

LEMMA 6.13. *Let $\mathbb{N}$ be the set of nonnegative integers. Let $c_1, c_2, \ldots, c_r$ be a family of cycles passing through a common state i. Assume that c_k and c_l are similar for some two indices $1 \le k, l \le r$. Then for any $n_1, n_2, \ldots, n_r \in \mathbb{N}$,*

$$\frac{P_i(N_t^{c_1} = n_1, \ldots, N_t^{c_k} = n_k, \ldots, N_t^{c_l} = n_l, \ldots, N_t^{c_r} = n_r)}{P_i(N_t^{c_1} = n_1, \ldots, N_t^{c_k} = n_l, \ldots, N_t^{c_l} = n_k, \ldots, N_t^{c_r} = n_r)} = \left(\frac{\gamma^{c_k}}{\gamma^{c_l}}\right)^{n_k - n_l}. \tag{93}$$

PROOF. We only need to prove this lemma when $k = 1$ and $l = 2$. The proof of the other cases is totally the same. To simplify notation, let $N = n_1 + \cdots + n_r$ and let

$$p = P_i\bigl(T^{c_1} < T^{c_2}\bigr), \qquad q = P_i\bigl(T^{c_2} < T^{c_1}\bigr). \tag{94}$$

Let T_n be the nth forming time of $c_1, c_2, \ldots, c_r$. Let $\tau_n = T_n - T_{n-1}$. Let ξ_n be the random variable defined in (76). Note that $\xi_n = c_k$ if and only if X forms c_k at time T_n. By distinguishing which one of $c_1, c_2, \ldots, c_r$ is formed at times $T_1, T_2, \ldots, T_N$, we obtain that

$$\begin{aligned} &P_i\bigl(N_t^{c_1} = n_1, \ldots, N_t^{c_r} = n_r\bigr) \\ &\quad = \sum_{A_1, \ldots, A_r} P_i(T_N \le t < T_{N+1}, \xi_m = c_1 \text{ for those } m \in A_1, \ldots, \xi_m = c_r \\ &\qquad \text{for those } m \in A_r), \end{aligned}$$

where the sequence of sets $A_1, \dots, A_r$ ranges over all partitions of $\{1, 2, \dots, N\}$ such that $\mathrm{Card}(A_k) = n_k$ for each $1 \le k \le r$. Then Lemma 6.12, together with the fact that $(\xi_n, \tau_n)_{n\ge1}$ is an i.i.d. sequence, shows that

$$
\begin{aligned}
&P_i\bigl(N_t^{c_1} = n_1, N_t^{c_2} = n_2, N_t^{c_3} = n_3, \dots, N_t^{c_r} = n_r\bigr)\\
&\quad = \sum_{A_1,\dots,A_r} P_i\bigl(T_N \le t < T_{N+1}, \xi_m \in \{c_1, c_2\} \text{ for those } m \in A_1 \cup A_2,\\
&\qquad\quad \xi_m = c_3 \text{ for those } m \in A_3, \dots, \xi_m = c_r \text{ for those } m \in A_r\bigr) p^{n_1} q^{n_2}\\
&\quad = \sum_{B_1,\dots,B_r} P_i\bigl(T_N \le t < T_{N+1}, \xi_m \in \{c_1, c_2\} \text{ for those } m \in B_2,\\
&\qquad\quad \xi_m = c_3 \text{ for those } m \in B_3, \dots, \xi_m = c_r \text{ for those } m \in B_r\bigr) C_{n_1+n_2}^{n_1} p^{n_1} q^{n_2}\\
&\quad = P_i\bigl(N_t^{c_1} + N_t^{c_2} = n_1 + n_2, N_t^{c_3} = n_3, \dots, N_t^{c_r} = n_r\bigr) C_{n_1+n_2}^{n_1} p^{n_1} q^{n_2},
\end{aligned}
$$

where the sequence of sets $B_2, \dots, B_r$ ranges over all partitions of $\{1, 2, \dots, N\}$ such that $\mathrm{Card}(B_2) = n_1 + n_2$ and $\mathrm{Card}(B_k) = n_k$ for each $3 \le k \le r$. By Theorem 5.5, it follows that

$$
\begin{aligned}
&P_i\bigl(N_t^{c_1} = n_1, N_t^{c_2} = n_2, N_t^{c_3} = n_3, \dots, N_t^{c_r} = n_r\bigr)\\
&\quad = P_i\bigl(N_t^{c_1} + N_t^{c_2} = n_1 + n_2, N_t^{c_3} = n_3, \dots, N_t^{c_r} = n_r\bigr) C_{n_1+n_2}^{n_1} p^{n_1} q^{n_2}\\
&\quad = P_i\bigl(N_t^{c_1} + N_t^{c_2} = n_1 + n_2, N_t^{c_3} = n_3, \dots, N_t^{c_r} = n_r\bigr)\\
&\qquad \times C_{n_1+n_2}^{n_1} p^{n_2} q^{n_1} \left(\frac{p}{q}\right)^{n_1-n_2}\\
&\quad = P_i\bigl(N_t^{c_1} = n_2, N_t^{c_2} = n_1, N_t^{c_3} = n_3, \dots, N_t^{c_r} = n_r\bigr) \left(\frac{\gamma^{c_1}}{\gamma^{c_2}}\right)^{n_1-n_2},
\end{aligned}
\tag{95}
$$

which gives the desired result. $\square$

The following lemma shows that the moment generating function of the sample circulations has a certain symmetry.

LEMMA 6.14. *Let $c_1, c_2, \dots, c_r$ be a family of cycles passing through a common state i. Assume that c_k and c_l are similar for some two indices $1 \le k, l \le r$. Let*

$$
g_t(\lambda_1, \dots, \lambda_r) = E_i e^{\lambda_1 N_t^{c_1} + \dots + \lambda_r N_t^{c_r}} = E_i e^{t(\lambda_1 J_t^{c_1} + \dots + \lambda_r J_t^{c_r})}.
\tag{96}
$$

Then for each $t \ge 0$ and any $\lambda_1, \dots, \lambda_r \in \mathbb{R}$,

$$
\begin{aligned}
&g_t(\lambda_1, \dots, \lambda_k, \dots, \lambda_l, \dots, \lambda_r)\\
&\quad = g_t\left(\lambda_1, \dots, \lambda_l - \log\frac{\gamma^{c_k}}{\gamma^{c_l}}, \dots, \lambda_k + \log\frac{\gamma^{c_k}}{\gamma^{c_l}}, \dots, \lambda_r\right).
\end{aligned}
\tag{97}
$$

PROOF. We only need to prove this lemma when $k = 1$ and $l = 2$. The proof of the other cases is totally the same. By Lemma 6.13, we have

$$
\begin{aligned}
&g_t(\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_r)\\
&\quad = E_i e^{\lambda_1 N_t^{c_1} + \lambda_2 N_t^{c_2} + \lambda_3 N_t^{c_3} + \cdots + \lambda_r N_t^{c_r}}\\
&\quad = \sum_{n_1, \dots, n_r \in \mathbb{N}} e^{\lambda_1 n_1 + \cdots + \lambda_r n_r} P_i\big(N_t^{c_1} = n_1, N_t^{c_2} = n_2, N_t^{c_3} = n_3, \dots, N_t^{c_r} = n_r\big)\\
&\quad = \sum_{n_1, \dots, n_r \in \mathbb{N}} e^{\lambda_1 n_1 + \cdots + \lambda_r n_r} P_i\big(N_t^{c_1} = n_2, N_t^{c_2} = n_1, N_t^{c_3} = n_3, \dots, N_t^{c_r} = n_r\big)\\
&\qquad \times \left(\frac{\gamma^{c_1}}{\gamma^{c_2}}\right)^{n_1 - n_2}\\
&\quad = \sum_{n_1, \dots, n_r \in \mathbb{N}} e^{(\lambda_1 + \log \frac{\gamma^{c_1}}{\gamma^{c_2}}) n_1 + (\lambda_2 - \log \frac{\gamma^{c_1}}{\gamma^{c_2}}) n_2 + \lambda_3 n_3 + \cdots + \lambda_r n_r}\\
&\qquad \times P_i\big(N_t^{c_1} = n_2, N_t^{c_2} = n_1, N_t^{c_3} = n_3, \dots, N_t^{c_r} = n_r\big)\\
&\quad = E_i e^{(\lambda_2 - \log \frac{\gamma^{c_1}}{\gamma^{c_2}}) N_t^{c_1} + (\lambda_1 + \log \frac{\gamma^{c_1}}{\gamma^{c_2}}) N_t^{c_2} \lambda_3 N_t^{c_3} + \cdots + \lambda_r N_t^{c_r}}\\
&\quad = g_t\left(\lambda_2 - \log \frac{\gamma^{c_1}}{\gamma^{c_2}}, \lambda_1 + \log \frac{\gamma^{c_1}}{\gamma^{c_2}}, \lambda_3, \dots, \lambda_r\right),
\end{aligned}
$$

which gives the desired result. □

The following result is a strengthened version of Varadhan's lemma in the large deviation theory.

LEMMA 6.15. *Let $(\mu_t)_{t>0}$ be a sequence of probability measures on a Polish space E which satisfies a large deviation principle with rate t and good rate function $I : E \to [0, \infty]$. Let $F : E \to \mathbb{R}$ be a continuous function. Assume that there exists $\gamma > 1$ such that the following moment condition is satisfied*:

$$
\limsup_{t \to \infty} \frac{1}{t} \log \int_E e^{\gamma t F(x)} \, d\mu_t(x) < \infty. \tag{98}
$$

Then

$$
\lim_{t \to \infty} \frac{1}{t} \log \int_E e^{t F(x)} \, d\mu_t(x) = \sup_{x \in E} \big(F(x) - I(x)\big). \tag{99}
$$

PROOF. The proof of this lemma can be found in [9], Theorem 4.3.1. □

Using the above lemma, we can obtain the following result.

LEMMA 6.16. *The notation is the same as in Lemma 6.14. Then for each* $\lambda \in \mathbb{R}^r$,

$$\lim_{t\to\infty} \frac{1}{t} \log g_t(\lambda) = \sup_{x\in\mathbb{R}^r} \{\lambda \cdot x - I^{c_1,c_2,\dots,c_r}(x)\}. \tag{100}$$

PROOF. Let $\lambda = (\lambda_1, \dots, \lambda_r)$. By Theorem 6.9, the law of $(J_t^{c_1}, \dots, J_t^{c_r})$ satisfies a large deviation principle with rate t and good rate function $I^{c_1,\dots,c_r}$. By Lemma 6.15, the result of this lemma holds if the following moment condition is satisfied for each $\gamma > 0$:

$$\limsup_{t\to\infty} \frac{1}{t} \log E_i e^{\gamma t(\lambda_1 J_t^{c_1} + \dots + \lambda_r J_t^{c_r})} < \infty. \tag{101}$$

Note that

$$\begin{aligned} E_i e^{\gamma t(\lambda_1 J_t^{c_1} + \dots + \lambda_r J_t^{c_r})} &\le E_i e^{\gamma |\lambda_1| N_t^{c_1} + \dots + \gamma |\lambda_r| N_t^{c_r}} \\ &\le E_i e^{\gamma\alpha(N_t^{c_1} + \dots + N_t^{c_r})} = E_i e^{\gamma\alpha N_t}, \end{aligned} \tag{102}$$

where $\alpha = \max\{|\lambda_1|, \dots, |\lambda_r|\}$ and $N_t = \inf\{n \ge 0 : T_{n+1} > t\}$ is the number of times that $c_1, \dots, c_r$ are formed by X up to time t. Since $X_0 = i$, in order to form any one of $c_1, \dots, c_r$, X must first leave state i. This shows that the nth forming time T_n of $c_1, \dots, c_r$ is larger than the sum of n independent exponential random variables with rate q_i, where $q_i = \sum_{j\neq i} q_{ij}$. This further implies that N_t is stochastically dominated by a Poisson random variable R_t with parameter $q_i t$. Thus, we obtain that

$$\begin{aligned} E_i e^{\gamma\alpha N_t} &= \int_{-\infty}^{\infty} \gamma\alpha e^{\gamma\alpha x} P_i(N_t \ge x)\,dx \le \int_{-\infty}^{\infty} \gamma\alpha e^{\gamma\alpha x} P_i(R_t \ge x)\,dx \\ &= E_i e^{\gamma\alpha R_t} = \sum_{n=0}^{\infty} e^{\gamma\alpha n} \frac{(q_i t)^n}{n!} e^{-q_i t} = \exp\bigl((e^{\gamma\alpha} - 1) q_i t\bigr). \end{aligned} \tag{103}$$

This shows that

$$\begin{aligned} \limsup_{t\to\infty} \frac{1}{t} \log E_i e^{\gamma t(\lambda_1 J_t^{c_1} + \dots + \lambda_r J_t^{c_r})} &\le \limsup_{t\to\infty} \frac{1}{t} \log E_i e^{\gamma\alpha N_t} \\ &\le \bigl(e^{\gamma\alpha} - 1\bigr) q_i < \infty. \end{aligned} \tag{104}$$

This completes the proof of this lemma. □

REMARK 6.17. Lemma 6.16 shows that

$$\lim_{t\to\infty} \frac{1}{t} \log g_t(\lambda) = \bigl(I^{c_1,c_2,\dots,c_r}\bigr)^*(\lambda), \tag{105}$$

where $(I^{c_1,c_2,\dots,c_r})^*$ is the Legendre–Fenchel transform of the rate function $I^{c_1,c_2,\dots,c_r}$. Recall that the Legendre–Fenchel transform of a function $f:\mathbb{R}^r\to[-\infty,\infty]$ is a function $f^*:\mathbb{R}^r\to[-\infty,\infty]$ defined by

$$f^*(\lambda)=\sup_{x\in\mathbb{R}^r}\{\lambda\cdot x-F(x)\}. \tag{106}$$

The following lemma, which is called the Fenchel–Moreau theorem, gives the sufficient and necessary conditions under which the Legendre–Fenchel transform is an involution. Recall that a function $f:\mathbb{R}^r\to[-\infty,\infty]$ is called proper if $f(x)<\infty$ for at least one x and $f(x)>-\infty$ for each x.

LEMMA 6.18. *Let $f:\mathbb{R}^r\to[-\infty,\infty]$ be a proper function. Then $f^{**}=f$ if and only if f is convex and lower semi-continuous, where $f^{**}=(f^*)^*$.*

PROOF. The proof of this lemma can be found in [6], Theorem 4.2.1. □

We are now in a position to prove the symmetry of the rate function $I^{c_1,c_2,\dots,c_r}$.

PROOF OF THEOREM 6.10. We only need to prove this theorem when $k=1$ and $l=2$. The proof of the other cases is totally the same. By Lemma 6.16, we have

$$\lim_{t\to\infty}\frac{1}{t}\log g_t(\lambda_1,\dots,\lambda_r)=\left(I^{c_1,\dots,c_r}\right)^*(\lambda_1,\dots,\lambda_r). \tag{107}$$

By Lemma 6.14, we have

$$g_t(\lambda_1,\lambda_2,\lambda_3,\dots,\lambda_r)=g_t\left(\lambda_2-\log\frac{\gamma^{c_1}}{\gamma^{c_2}},\lambda_1+\log\frac{\gamma^{c_1}}{\gamma^{c_2}},\lambda_3,\dots,\lambda_r\right). \tag{108}$$

Combining the above two equations, we obtain that

$$\begin{aligned}&\left(I^{c_1,\dots,c_r}\right)^*(\lambda_1,\lambda_2,\lambda_3,\dots,\lambda_r)\\&\qquad=\left(I^{c_1,\dots,c_r}\right)^*\left(\lambda_2-\log\frac{\gamma^{c_1}}{\gamma^{c_2}},\lambda_1+\log\frac{\gamma^{c_1}}{\gamma^{c_2}},\lambda_3,\dots,\lambda_r\right).\end{aligned} \tag{109}$$

By Theorem 6.9 and Lemma 6.11, $I^{c_1,\dots,c_r}$ is a good rate function which is also convex. This shows that $I^{c_1,\dots,c_r}$ is proper, convex and lower semicontinuous. By Lemma 6.18, we obtain that $I^{c_1,\dots,c_r}=(I^{c_1,\dots,c_r})^{**}$. Thus, we have

$$\begin{aligned}&I^{c_1,\dots,c_r}(x_1,x_2,x_3,\dots,x_r)\\&\quad=\left(I^{c_1,\dots,c_r}\right)^{**}(x_1,x_2,x_3,\dots,x_r)\\&\quad=\sup_{\lambda_1,\dots,\lambda_r\in\mathbb{R}}\left\{\lambda_1x_1+\cdots+\lambda_rx_r-\left(I^{c_1,\dots,c_r}\right)^*(\lambda_1,\lambda_2,\lambda_3,\dots,\lambda_r)\right\}\\&\quad=\sup_{\lambda_1,\dots,\lambda_r\in\mathbb{R}}\Bigg\{\lambda_1x_1+\cdots+\lambda_rx_r\end{aligned}$$

$$
\begin{aligned}
&\quad - \left(I^{c_1,\dots,c_r}\right)^*\left(\lambda_2 - \log\frac{\gamma^{c_1}}{\gamma^{c_2}}, \lambda_1 + \log\frac{\gamma^{c_1}}{\gamma^{c_2}}, \lambda_3, \dots, \lambda_r\right)\Bigg\} \\
&= \sup_{\lambda_1,\dots,\lambda_r\in\mathbb{R}} \Bigg\{\left(\lambda_1 - \log\frac{\gamma^{c_1}}{\gamma^{c_2}}\right)x_1 + \left(\lambda_2 + \log\frac{\gamma^{c_1}}{\gamma^{c_2}}\right)x_2 + \lambda_3 x_3 + \dots + \lambda_r x_r \\
&\quad - \left(I^{c_1,\dots,c_r}\right)^*(\lambda_2, \lambda_1, \lambda_3, \dots, \lambda_r)\Bigg\} \\
&= I^{c_1,\dots,c_r}(x_2, x_1, x_3, \dots, x_r) - \left(\log\frac{\gamma^{c_1}}{\gamma^{c_2}}\right)(x_1 - x_2),
\end{aligned}
$$

which gives the desired result. □

REMARK 6.19. Let

$$
g(\lambda) = \lim_{t\to\infty}\frac{1}{t}\log g_t(\lambda) = \lim_{t\to\infty}\frac{1}{t}\log E_i e^{t(\lambda_1 J_t^{c_1} + \dots + \lambda_r J_t^{c_r})}. \tag{110}
$$

The above proof shows that

$$
I^{c_1,\dots,c_r}(x) = g^*(x) = \sup_{\lambda\in\mathbb{R}^r}\{\lambda\cdot x - g(\lambda)\}. \tag{111}
$$

This equality is closely related to the well-known Gartner–Ellis theorem in the large deviation theory. Roughly speaking, the Gartner–Ellis theorem claims that if $g(\lambda)$ is differentiable everywhere, then the sample circulations $J_t^{c_1}, J_t^{c_2}, \dots, J_t^{c_r}$ must satisfy a large deviation principle with a good rate function which has the form of (111). This is the routine approach to study the large deviations and fluctuation theorems in nonequilibrium statistical physics. However, in some complicated problems, it is extremely difficult to prove the differentiability of $g(\lambda)$. Therefore, in this paper, we obtain the desired results with the aid of the large deviations of Markov renewal processes without involving into the complex differentiability issue.

Since a cycle c and its reversed cycle $c-$ must be similar, we obtain the following corollary.

COROLLARY 6.20. *Let $c_1, c_2, \dots, c_r$ be a family of cycles passing through a common state i. Then under P_i, the law of $(J_t^{c_1}, J_t^{c_1-}, \dots, J_t^{c_r}, J_t^{c_r-})$ satisfies a large deviation principle with rate t and good rate function $I^{c_1,c_1-,\dots,c_r,c_r-}: \mathbb{R}^{2r} \to [0,\infty]$. Moreover, for each $1 \le k \le r$, the rate function $I^{c_1,c_1-,\dots,c_r,c_r-}$ has the following symmetry: for any $x_1, y_1, \cdots, x_r, y_r \in \mathbb{R}$ and $1 \le k \le r$,*

$$
\begin{aligned}
&I^{c_1,c_1-,\dots,c_r,c_r-}(x_1, y_1, \dots, x_k, y_k, \dots, x_r, y_r) \\
&\quad = I^{c_1,c_1-,\dots,c_r,c_r-}(x_1, y_1, \dots, y_k, x_k, \dots, x_r, y_r) \\
&\qquad - \left(\log\frac{\gamma^{c_k}}{\gamma^{c_k-}}\right)(x_k - y_k).
\end{aligned} \tag{112}
$$

PROOF. This corollary follows directly from Theorems 6.9 and 6.10. □

REMARK 6.21. The generalized Haldane equalities characterize the symmetry of the forming times of a family of similar cycles. Theorem 6.10, Lemma 6.13 and Lemma 6.14, however, characterize the symmetry of the sample circulations along a family of similar cycles from different aspects.

7. Applications in natural sciences.

7.1. *Applications in nonequilibrium statistical physics.* Markov chains are widely used to model various stochastic systems in physics, chemistry and biology. In nonequilibrium statistical physics, one of the most important concepts associated with a stochastic system is the entropy production rate. Let $X = (X_t)_{t\geq 0}$ be an irreducible and recurrent continuous-time Markov chain with denumerable state space S and transition rate matrix $Q = (q_{ij})$. Then the sample entropy production rate W_t of X by time t is defined as

$$\begin{aligned} W_t &= \frac{1}{t}\log\frac{p_0(X_0)q_{\bar{X}_0\bar{X}_1}q_{\bar{X}_1\bar{X}_2}\cdots q_{\bar{X}_{\tilde{N}_t-1}\bar{X}_{\tilde{N}_t}}}{p_t(X_t)q_{\bar{X}_1\bar{X}_0}q_{\bar{X}_2\bar{X}_1}\cdots q_{\bar{X}_{\tilde{N}_t}\bar{X}_{\tilde{N}_t-1}}} \\ &= \frac{1}{t}\log\frac{p_0(X_0)}{p_t(X_t)} + \frac{1}{t}\sum_{i=0}^{\tilde{N}_t-1}\log\frac{q_{\bar{X}_i\bar{X}_{i+1}}}{q_{\bar{X}_{i+1}\bar{X}_i}}, \end{aligned} \tag{113}$$

where $p_t = (p_t(i))$ is the probability distribution of X at time t, $\bar{X} = (\bar{X}_n)_{n\geq 0}$ is the embedded chain of X, and $\tilde{N}_t$ is the number of jumps of X by time t. In recent years, physicists found that the sample entropy production rate W_t satisfies various types of fluctuation theorems [22, 28, 42]. This discovery has been considered as one of the most important results in nonequilibrium statistical physics in the last two decades.

Interestingly, the sample entropy production rate of Markov chains can be decomposed along different cycles [21, 36, 43]. Specifically, the sample entropy production rate W_t can be decomposed as

$$W_t = \frac{1}{2}\sum_c K_t^c \log\frac{\gamma^c}{\gamma^{c-}} + W_t^r, \tag{114}$$

where c ranges over all cycles, K_t^c is the sample net circulation along cycle c, γ^c is the strength of cycle c, and the remainder W_t^r collects the contributions of those state transitions by time t that do not form a full cycle. This implies that the sample net circulation K_t^c along cycle c is proportional to the sample entropy production rate of X along cycle c. Thus, it is nature to ask whether we can establish the fluctuation theorems for the sample net circulations of Markov chains.

Fortunately, the generalized Haldane equalities established in this paper can be used to study the fluctuation theorems for the sample net circulations. To make the readers understand the connections between our work and nonequilibrium statistical physics, we briefly state various types of fluctuation theorems for the sample net circulations.

The following definition of the affinity originates from nonequilibrium statistical physics [43].

DEFINITION 7.1. Let c be a cycle. Then the affinity of cycle c is defined as

$$\rho^c = \log \frac{\gamma^c}{\gamma^{c-}}. \tag{115}$$

Theorems of the following type are called transient fluctuation theorems in nonequilibrium statistical physics.

THEOREM 7.2. *Let $c_1, c_2, \ldots, c_r$ be a family of cycles passing through a common state i. Then for any $n_1, n_2, \ldots, n_r \in \mathbb{Z}$ and $1 \le k \le r$,*

$$\frac{P_i(K_t^{c_1} = n_1/t, \ldots, K_t^{c_k} = n_k/t, \ldots, K_t^{c_r} = n_r/t)}{P_i(K_t^{c_1} = n_1/t, \ldots, K_t^{c_k} = -n_k/t, \ldots, K_t^{c_r} = n_r/t)} = e^{n_k \rho^{c_k}}. \tag{116}$$

PROOF. We only need to prove this theorem when $k = 1$. The proof of the other cases is totally the same. By Lemma 6.13, we have

$$\begin{aligned}
&P_i\big(K_t^{c_1} = n_1/t, \ldots, K_t^{c_k} = n_k/t, \ldots, K_t^{c_r} = n_r/t\big)\\
&\quad = P_i\big(N_t^{c_1} - N_t^{c_1-} = n_1, N_t^{c_2} - N_t^{c_2-} = n_2, \ldots, N_t^{c_r} - N_t^{c_r-} = n_r\big)\\
&\quad = \sum_{l_1-m_1=n_1,\ldots,l_r-m_r=n_r} P_i\big(N_t^{c_1} = l_1, N_t^{c_1-} = m_1, N_t^{c_2} = l_2, N_t^{c_2-} = m_2, \ldots,\\
&\qquad N_t^{c_r} = l_r, N_t^{c_r-} = m_r\big)\\
&\quad = \sum_{l_1-m_1=n_1,\ldots,l_r-m_r=n_r} P_i\big(N_t^{c_1} = m_1, N_t^{c_1-} = l_1, N_t^{c_2} = l_2, N_t^{c_2-} = m_2, \ldots,\\
&\qquad N_t^{c_r} = l_r, N_t^{c_r-} = m_r\big)e^{(l_1-m_1)\rho^{c_1}}\\
&\quad = \sum_{l_1-m_1=-n_1,\ldots,l_r-m_r=n_r} P_i\big(N_t^{c_1} = l_1, N_t^{c_1-} = m_1, N_t^{c_2} = l_2, N_t^{c_2-} = m_2, \ldots,\\
&\qquad N_t^{c_r} = l_r, N_t^{c_r-} = m_r\big)e^{n_1\rho^{c_1}}\\
&\quad = P_i\big(N_t^{c_1} - N_t^{c_1-} = -n_1, N_t^{c_2} - N_t^{c_2-} = n_2, \ldots, N_t^{c_r} - N_t^{c_r-} = n_r\big)e^{n_1\rho^{c_1}}\\
&\quad = P_i\big(K_t^{c_1} = n_1/t, \ldots, K_t^{c_k} = -n_k/t, \ldots, K_t^{c_r} = n_r/t\big)e^{n_1\rho^{c_1}},
\end{aligned}$$

which gives the desired result. □

Theorems of the following type are called Kurchan–Lebowitz–Spohn-type fluctuation theorems in nonequilibrium statistical physics.

THEOREM 7.3. *Let $c_1, c_2, \ldots, c_r$ be a family of cycles passing through a common state i. Let*

$$h_t(\lambda_1, \ldots, \lambda_r) = E_i e^{t(\lambda_1 K_t^{c_1} + \cdots + \lambda_r K_t^{c_r})}. \tag{117}$$

Then for any $t \geq 0$, $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$, and $1 \leq k \leq r$,

$$h_t(\lambda_1, \ldots, \lambda_k, \ldots, \lambda_r) = h_t\bigl(\lambda_1, \ldots, -(\lambda_k + \rho^{c_k}), \ldots, \lambda_r\bigr). \tag{118}$$

PROOF. We only need to prove this theorem when $k = 1$. The proof of the other cases is totally the same. By Theorem 7.2, we have

$$\begin{aligned}
h_t&(\lambda_1, \lambda_2, \ldots, \lambda_r)\\
&= E_i e^{t(\lambda_1 K_t^{c_1} + \cdots + \lambda_r K_t^{c_r})}\\
&= \sum_{n_1, \ldots, n_r \in \mathbb{Z}} e^{\lambda_1 n_1 + \lambda_2 n_2 + \cdots + \lambda_r n_r}\\
&\qquad \times P_i\bigl(K_t^{c_1} = n_1/t, K_t^{c_2} = n_2/t, \ldots, K_t^{c_r} = n_r/t\bigr)\\
&= \sum_{n_1, \ldots, n_r \in \mathbb{Z}} e^{\lambda_1 n_1 + \lambda_2 n_2 + \cdots + \lambda_r n_r}\\
&\qquad \times P_i\bigl(K_t^{c_1} = -n_1/t, K_t^{c_2} = n_2/t, \ldots, K_t^{c_r} = n_r/t\bigr) e^{n_1 \rho^{c_1}}\\
&= \sum_{n_1, \ldots, n_r \in \mathbb{Z}} e^{(\lambda_1 + \rho^{c_1}) n_1 + \lambda_2 n_2 + \cdots + \lambda_r n_r}\\
&\qquad \times P_i\bigl(K_t^{c_1} = -n_1/t, K_t^{c_2} = n_2/t, \ldots, K_t^{c_r} = n_r/t\bigr)\\
&= E_i e^{t(-(\lambda_1 + \rho^{c_1}) K_t^{c_1} + \lambda_2 K_t^{c_2} + \cdots + \lambda_r K_t^{c_r})} = h_t\bigl(-(\lambda_1 + \rho^{c_1}), \lambda_2, \ldots, \lambda_r\bigr),
\end{aligned}$$

which gives the desired result. □

Theorems of the following type are called integral fluctuation theorems in nonequilibrium statistical physics.

THEOREM 7.4. *Let $c_1, c_2, \ldots, c_r$ be a family of cycles passing through a common state i. Then for each $t \geq 0$,*

$$E_i e^{-t(K_t^{c_1} \rho^{c_1} + K_t^{c_2} \rho^{c_2} + \cdots + K_t^{c_r} \rho^{c_r})} = 1. \tag{119}$$

PROOF. By Theorem 7.3, for any $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$,

$$E_i e^{t(\lambda_1 K_t^{c_1} + \cdots + \lambda_r K_t^{c_r})} = E_i e^{-t((\lambda_1 + \rho^{c_1}) K_t^{c_1} + \cdots + (\lambda_r + \rho^{c_r}) K_t^{c_r})}. \tag{120}$$

If we take $\lambda_k = -\rho^{c_k}$ for each k in the above equation, we obtain the desired result. □

The large deviation principle of the sample net circulations and the symmetry of the rate function are stated in the following theorem. Theorems of the following type are called Gallavotti–Cohen-type fluctuation theorems in nonequilibrium statistical physics.

THEOREM 7.5. *Let $c_1, c_2, \ldots, c_r$ be a family of cycles passing through a common state i. Then under P_i, the law of $(K_t^{c_1}, K_t^{c_2}, \ldots, K_t^{c_r})$ satisfies a large deviation principle with rate t and good rate function $I_K^{c_1,c_2,\ldots,c_r} : \mathbb{R}^r \to [0,\infty]$. Moreover, the rate function $I_K^{c_1,c_2,\ldots,c_r}$ has the following symmetry: for any $x_1, x_2, \ldots, x_r \in \mathbb{R}$ and $1 \le k \le r$,*

$$I_K^{c_1,c_2,\ldots,c_r}(x_1,\ldots,x_k,\ldots,x_r) = I_K^{c_1,c_2,\ldots,c_r}(x_1,\ldots,-x_k,\ldots,x_r) - \rho^{c_k}x_k. \tag{121}$$

PROOF. We only need to prove this theorem when $k = 1$. The proof of the other cases is totally the same. Let $F : \mathbb{R}^{2r} \to \mathbb{R}^r$ be a continuous map defined as

$$F(x_1, y_1, \ldots, x_r, y_r) = (x_1 - y_1, \ldots, x_r - y_r). \tag{122}$$

Then we have

$$F\big(J_t^{c_1}, J_t^{c_1-}, \ldots, J_t^{c_r}, J_t^{c_r-}\big) = \big(K_t^{c_1}, \ldots, K_t^{c_r}\big). \tag{123}$$

By Corollary 6.20, the law of $(J_t^{c_1}, J_t^{c_1-}, \ldots, J_t^{c_r}, J_t^{c_r-})$ satisfies a large deviation principle with rate t and good rate function $I^{c_1,c_1-,\ldots,c_r,c_r-}$. Using the contraction principle, we see that the law of $(K_t^{c_1}, \ldots, K_t^{c_r})$ satisfies a large deviation principle with rate t and good rate function

$$\begin{aligned} &I_K^{c_1,\ldots,c_r}(z_1,\ldots,z_r) \\ &\qquad = \inf_{x_1-y_1=z_1,\ldots,x_r-y_r=z_r} I^{c_1,c_1-,\ldots,c_r,c_r-}(x_1,y_1,\ldots,x_r,y_r). \end{aligned} \tag{124}$$

Thus, we have

$$\begin{aligned} I_K^{c_1,\ldots,c_r}&(z_1,z_2,\ldots,z_r) \\ &= \inf_{x_1-y_1=z_1,\ldots,x_r-y_r=z_r} I^{c_1,c_1-,\ldots,c_r,c_r-}(x_1,y_1,x_2,y_2,\ldots,x_r,y_r) \\ &= \inf_{x_1-y_1=z_1,\ldots,x_r-y_r=z_r} I^{c_1,c_1-,\ldots,c_r,c_r-}(y_1,x_1,x_2,y_2,\ldots,x_r,y_r) \\ &\quad - \rho^{c_1}(x_1-y_1) \\ &= \inf_{y_1-x_1=-z_1,x_2-y_2=z_2,\ldots,x_r-y_r=z_r} I^{c_1,c_1-,\ldots,c_r,c_r-}(y_1,x_1,x_2,y_2,\ldots,x_r,y_r) \\ &\quad - \rho^{c_1}z_1 \\ &= I_K^{c_1,\ldots,c_r}(-z_1,z_2,\ldots,z_r) - \rho^{c_1}z_1, \end{aligned}$$

which gives the desired result. $\square$

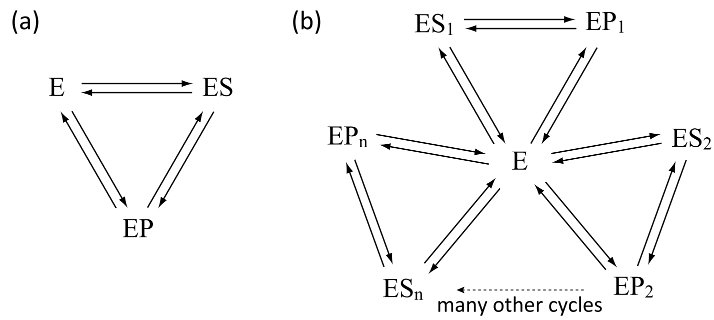

FIG. 1. *Markov chain models of enzyme kinetics*. (a) *The Markov chain model of single-substrate enzyme kinetics*. (b) *The Markov chain model of multiple-substrate enzyme kinetics*.

7.2. *Applications in biochemistry.* One of the most important branches of biochemistry is enzyme kinetics, which studies chemical reactions catalyzed by enzymes. In recent years, it has been made possible to study enzyme kinetics at the single-molecule level [11, 18, 32], in which case the concept of concentration makes no sense and the behavior of enzymes must be studied in a single-molecule way.

Let us consider the following three-step Michaelis–Menten enzyme kinetics [5, 16, 18]:

$$E + S \rightleftharpoons ES \rightleftharpoons EP \rightleftharpoons E + P, \tag{125}$$

where E is an enzyme turning the substrate S into the product P. If there is only one enzyme molecule, then it may transition stochastically among three states: the free enzyme E, the enzyme-substrate complex ES and the enzyme-product complex EP. From the perspective of a single enzyme molecule, the Michaelis–Menten enzyme kinetics (125) can be modeled as a three-state Markov chain illustrated in Figure 1(a).

However, single-substrate enzymes are actually rather rare in biochemistry [17]. If the enzyme E can catalyze multiple chemical reactions simultaneously with substrates $S_1, S_2, \ldots, S_n$ and products $P_1, P_2, \ldots, P_n$, then the transition diagram of the Markov chain model will contain multiple cycles passing through a common state E, as illustrated in Figure 1(b).

We assume that the Markov chain illustrated in Figure 1(b) starts from state E. If the Markov chain forms a clockwise cycle $c_k = (E, ES_k, EP_k)$, then the substrate S_k is converted into the product P_k for one time. Similarly, if the Markov chain forms a counterclockwise cycle $c_k- = (E, EP_k, ES_k)$, then the product P_k is converted into the substrate S_k for one time. Thus, the sample net circulation $K_t^{c_k}$ along cycle c_k represents the net number of conversions from the substrate S_k into the product P_k per unit time and the quantity

$$W_t^{c_k} = K_t^{c_k} \rho^{c_k} \tag{126}$$

represents the fluctuating chemical work done along cycle c_k [17, 18, 33], where ρ^{c_k} is the affinity of cycle c_k. In fact, the results of this paper can be directly applied to establish the multivariate fluctuation theorems for the sample net circulations along cycles $c_1, c_2, \ldots, c_n$ and the fluctuating chemical works done along cycles $c_1, c_2, \ldots, c_n$. This shows that our work could have a broad application prospect in biochemistry.

Acknowledgments. The authors gratefully acknowledge H. Ge at Peking University and H. Qian at the University of Washington for stimulating discussions and the anonymous reviewers for valuable suggestions.

REFERENCES

[1] ANDRIEUX, D. and GASPARD, P. (2007). Fluctuation theorem for currents and Schnakenberg network theory. *J. Stat. Phys.* **127** 107–131. MR2313063

[2] ANDRIEUX, D. and GASPARD, P. (2007). Network and thermodynamic conditions for a single macroscopic current fluctuation theorem. *C. R. Phys.* **8** 579–590.

[3] BARATO, A. C. and CHETRITE, R. (2012). On the symmetry of current probability distributions in jump processes. *J. Phys. A* **45** 485002, 23. MR2998415

[4] BAUER, M. and CORNU, F. (2014). Affinity and fluctuations in a mesoscopic noria. *J. Stat. Phys.* **155** 703–736. MR3192181

[5] BEARD, D. A. and QIAN, H. (2008). *Chemical Biophysics*: *Quantitative Analysis of Cellular Systems*. Cambridge Univ. Press, Cambridge. MR2542822

[6] BORWEIN, J. M. and LEWIS, A. S. (2006). *Convex Analysis and Nonlinear Optimization*: *Theory and Examples*, 2nd ed. Springer, New York. MR2184742

[7] CHUNG, L. K. (1967). *Markov Chains*. Springer, Berlin.

[8] CROOKS, G. E. (1999). Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E* **60** 2721–2726.

[9] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. Springer, New York. MR1619036

[10] DUBINS, L. E. (1996). The gambler's ruin problem for periodic walks. In *Statistics*, *Probability and Game Theory*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **30** 7–12. IMS, Hayward, CA. MR1481769

[11] ENGLISH, B. P., MIN, W., VAN OIJEN, A. M., LEE, T. K., LUO, G., SUN, H., CHERAYIL, B. J., KOU, S. C. and XIE, X. S. (2005). Ever-fluctuating single enzyme molecules: Michaelis–Menten equation revisited. *Nat. Chem. Biol.* **2** 87–94.

[12] ESPOSITO, M. and VAN DEN BROECK, C. (2010). Three detailed fluctuation theorems. *Phys. Rev. Lett.* **104** 090601, 4. MR2608648

[13] EVANS, D. J., COHEN, E. G. D. and MORRISS, G. P. (1993). Probability of second law violations in shearing steady states. *Phys. Rev. Lett.* **71** 2401–2404.

[14] FAGGIONATO, A. and DI PIETRO, D. (2011). Gallavotti–Cohen-type symmetry related to cycle decompositions for Markov chains and biochemical applications. *J. Stat. Phys.* **143** 11–32. MR2787971

[15] GALLAVOTTI, G. and COHEN, E. G. D. (1995). Dynamical ensembles in stationary states. *J. Stat. Phys.* **80** 931–970. MR1349772

[16] GE, H. (2008). Waiting cycle times and generalized Haldane equality in the steady-state cycle kinetics of single enzymes. *J. Phys. Chem. B* **112** 61–70.

[17] GE, H. (2012). Multivariable fluctuation theorems in the steady-state cycle kinetics of single enzyme with competing substrates. *J. Phys. A*: *Math. Theor.* **45** 215002.

[18] GE, H., QIAN, M. and QIAN, H. (2012). Stochastic theory of nonequilibrium steady states. Part II: Applications in chemical biophysics. *Phys. Rep.* **510** 87–118.

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

[20] JARZYNSKI, C. (1997). Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **78** 2690–2693.

[21] JIANG, D.-Q., QIAN, M. and QIAN, M.-P. (2004). *Mathematical Theory of Nonequilibrium Steady States: On the Frontier of Probability and Dynamical Systems*. Springer, Berlin. MR2034774

[22] JIANG, D.-Q., QIAN, M. and ZHANG, F.-X. (2003). Entropy production fluctuations of finite Markov chains. *J. Math. Phys.* **44** 4176–4188. MR2003952

[23] KALPAZIDOU, S. (1990). Asymptotic behaviour of sample weighted circuits representing recurrent Markov chains. *J. Appl. Probab.* **27** 545–556. MR1067021

[24] KALPAZIDOU, S. L. (1995). *Cycle Representations of Markov Processes*. Springer, New York. MR1336140

[25] KOLMOGOROFF, A. (1936). Zur Theorie der Markoffschen Ketten. *Math. Ann.* **112** 155–160. MR1513044

[26] KOLOMEISKY, A. B., STUKALIN, E. B. and POPOV, A. A. (2005). Understanding mechanochemical coupling in kinesins using first-passage-time processes. *Phys. Rev. E* **71** 031902.

[27] KURCHAN, J. (1998). Fluctuation theorem for stochastic dynamics. *J. Phys. A* **31** 3719–3729. MR1630333

[28] LEBOWITZ, J. L. and SPOHN, H. (1999). A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics. *J. Stat. Phys.* **95** 333–365. MR1705590

[29] MARIANI, M. and ZAMBOTTI, L. (2014). Large deviations for the empirical measure of heavy tailed Markov renewal processes. Available at arXiv:1203.5930v2.

[30] POLETTINI, M. and ESPOSITO, M. (2014). Transient fluctuation theorems for the currents and initial equilibrium ensembles. *J. Stat. Mech. Theory Exp.* **10** P10033, 18. MR3278063

[31] QIAN, C., QIAN, M. and QIAN, M. P. (1981). Markov chain as a model of Hill's theory on circulation. *Sci. Sinica* **24** 1431–1448. MR0655943

[32] QIAN, H. and ELSON, E. L. (2002). Single-molecule enzymology: Stochastic Michaelis-Menten kinetics. *Biophys. Chem.* **101** 565–576.

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

[34] QIAN, M. P. and QIAN, M. (1979). The decomposition into a detailed balance part and a circulation part of an irreversible stationary Markov chain. *Sci. Sinica* 69–79. Special Issue II on Math. MR0576348

[35] QIAN, M. P. and QIAN, M. (1982). Circulation for recurrent Markov chains. *Z. Wahrsch. Verw. Gebiete* **59** 203–210. MR0650612

[36] QIAN, M. P., QIAN, M. and GONG, G. L. (1991). The reversibility and the entropy production of Markov processes. In *Probability Theory and Its Applications in China. Contemp. Math.* **118** 255–261. Amer. Math. Soc., Providence, RI. MR1137974

[37] QIAN, M. P., QIAN, M. and QIAN, C. (1982). Circulation distribution of a Markov chain—cycle skipping rate and decomposition according to probability meaning. *Sci. Sinica Ser. A* **25** 31–40. MR0669671

[38] QIAN, M. P., QIAN, M. and QIAN, C. (1984). Circulations of Markov chains with continuous time and the probability interpretation of some determinants. *Sci. Sinica Ser. A* **27** 470–481. MR0764734

[39] SAMUELS, S. M. (1975). The classical ruin problem with equal initial fortunes. *Math. Mag.* **48** 286–288. MR0413298

[40] SCHNAKENBERG, J. (1976). Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Modern Phys*. **48** 571–585. MR0443796

[41] SEGEL, I. H. (1975). *Enzyme Kinetics*. Wiley, New York.

[42] SEIFERT, U. (2005). Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Phys*. *Rev. Lett*. **95** 040602.

[43] SEIFERT, U. (2012). Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep*. *Progr. Phys*. **75** 126001.

[44] SPINNEY, R. E. and FORD, I. J. (2012). Nonequilibrium thermodynamics of stochastic systems with odd and even variables. *Phys*. *Rev. Lett*. **108** 170603.

[45] ZHANG, X.-J., QIAN, H. and QIAN, M. (2012). Stochastic theory of nonequilibrium steady states and its applications. Part I. *Phys*. *Rep*. **510** 1–86. MR2868686

C. JIA
BEIJING COMPUTATIONAL SCIENCE RESEARCH CENTER
ZHONGGUANCUN SOFTWARE PARK II,
NO. 10 WEST DONGBEIWANG ROAD
HAIDIAN DISTRICT, BEIJING 100094
P.R. CHINA
AND
LMAM
SCHOOL OF MATHEMATICAL SCIENCES
PEKING UNIVERSITY
NO. 5 YIHEYUAN ROAD
HAIDIAN DISTRICT, BEIJING 100871
P.R. CHINA
E-MAIL: jiac@pku.edu.cn

D.-Q. JIANG
LMAM
SCHOOL OF MATHEMATICAL SCIENCES
PEKING UNIVERSITY
NO. 5 YIHEYUAN ROAD
HAIDIAN DISTRICT, BEIJING 100871
P.R. CHINA
AND
CENTER FOR STATISTICAL SCIENCE
PEKING UNIVERSITY
NO. 5 YIHEYUAN ROAD
HAIDIAN DISTRICT, BEIJING 100871
P.R. CHINA
E-MAIL: jiangdq@math.pku.edu.cn

M.-P. QIAN
LMAM
SCHOOL OF MATHEMATICAL SCIENCES
PEKING UNIVERSITY
NO. 5 YIHEYUAN ROAD
HAIDIAN DISTRICT, BEIJING 100871
P.R. CHINA
E-MAIL: qianmp@math.pku.edu.cn