The transient fluctuation theorem of sample entropy production for general stochastic processes

2007 J. Phys. A: Math. Theor. 40 F713

(http://iopscience.iop.org/1751-8121/40/30/F04)

IOP FTC
IOP Publishing — Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **40** (2007) F713–F723 doi:10.1088/1751-8113/40/30/F04

FAST TRACK COMMUNICATION

The transient fluctuation theorem of sample entropy production for general stochastic processes

Hao Ge and Da-Quan Jiang

LMAM, School of Mathematical Sciences, Peking University, Beijing 100871, People's Republic of China

E-mail: edmund-ge@tom.com and jiangdq@math.pku.edu.cn

Received 15 February 2007, in final form 10 June 2007
Published 12 July 2007
Online at stacks.iop.org/JPhysA/40/F713

Abstract

The transient fluctuation theorem for stochastic processes was first put forward by D J Searles and D J Evans (1999). In the present paper, it is rigorously proved that the transient fluctuation theorem (TFT) of sample entropy production, which is previously defined by Jiang *et al* (2003) and Reid *et al* (2005) and also called the general action functional up to boundary terms by Lebowitz and Spohn (1999), holds for general stochastic processes without the assumption of Markovian, homogeneous or stationary properties. Then the condition of our theorem is verified for various stochastic processes, including homogeneous, inhomogeneous Markov chains and general diffusion processes. Among these cases, the transient fluctuation theorems for inhomogeneous Markov chains and general diffusion processes are rigorously derived for the first time.

PACS numbers: 05.40.−a, 05.70.−a, 05.70.Ln

1. Introduction

The fluctuation theorem (FT) gives a general formula valid in nonequilibrium systems, for the logarithm of the probability ratio of observing trajectories that satisfy or 'violate' the second law of thermodynamics. In 1993, Evans, Cohen and Morriss [6] found in computer simulations that the natural invariant measure of a stationary nonequilibrium system has a symmetry, which is later called the fluctuation theorem by Gallavotti and Cohen [9, 10]. Motivated by the result in [6], Gallavotti and Cohen [8, 10] gave the first mathematical presentation of the fluctuation theorem for stationary nonequilibrium systems modelled by hyperbolic dynamical systems: provided that the dynamics satisfies time reversal invariance and is sufficiently chaotic, the probability distributions of the phase space contraction averaged over a large time span t have a highly non-obvious symmetry, whose specific form is given by the large deviation rate function. Besides the steady-state fluctuation theorem, Evans and Searles [1, 7, 36–38] considered transient, nonequilibrium systems and employed a known equilibrium state as the initial distribution to derive a transient fluctuation theorem.

Except for the deterministic nonequilibrium systems, Kurchan [22] pointed out that the fluctuation theorem also holds for certain diffusion processes. Lebowitz and Spohn [23] extended Kurchan's results to quite general Markov processes, and investigated various examples. Furthermore, they mentioned that the logarithm of the Radon–Nikodym derivative of the process with respect to its time-reversal can be regarded to be a definition of the action functional in a general stochastic dynamics without Markovian condition. Searles and Evans [35] derived informally the transient fluctuation theorem for a class of non-stationary stochastic systems. However, their proof is not purely mathematically rigorous. In 2003, Jiang and Zhang [20] gave some rigorous mathematical results about the entropy production fluctuation for Markov chains, including both steady-state fluctuation theorem and transient fluctuation theorem.

Since the emergence of fluctuation theorems in 1993, their degree of universality is an interesting subject of investigation. It is therefore useful to have rigorous derivations for as much physical situations as possible in order to rule out the existence of counterexamples in particular physical systems.

In the present paper, we only focus on the derivation of the Evans–Searles fluctuation theorem [35] for general stochastic processes. Our definition of sample entropy production comes from the previous works [20] and [32]. It is also called the general action functional by Lebowitz and Spohn [23] up to boundary terms. We rigorously proved that the transient fluctuation theorem (TFT) of sample entropy production holds for general stochastic processes without the assumption of Markovian, homogeneous or stationary properties. Most of the previous works on this subject are not purely mathematically rigorous, or based on several additional assumptions. Therefore, although the derivation of theorem 2.2 is easy from the mathematical point of view, it is a unified result without any additional assumption on the underlying dynamics.

The concept of entropy production was first put forward in nonequilibrium statistical physics to describe how far a specific state of a system is from its equilibrium state [14, 25, 34]. In [28–30], a measure-theoretic definition of entropy production rate is given for stochastic processes, unifying different entropy production formulae in various concrete cases. Entropy production is defined as the specific relative entropy of the process with respect to its time reversal, which is just the time-averaged expectation of the action functional mentioned in [23]. Here, we recommend a recent book [19] on the mathematical theory of nonequilibrium steady states.

Recently, inhomogeneous stochastic processes have attracted much interest from statistical physicists, including the diffusion approximation for master equations [21] and the relationship between Jarzynski's equality and fluctuation theorems [2–4, 15–18].

Time inhomogeneity causes many difficulties in studying mathematically the physical properties of stochastic processes. We have investigated the statistical physical properties of inhomogeneous Markov chains since 2004, including reversibility, entropy production and generalized Jarzynski's equality [11–13].

We will verify the condition of our main result, theorem 2.2, for various stochastic processes in section 3, including homogeneous, inhomogeneous Markov chains and general diffusion processes. Among these cases, the applications of theorem 2.2 to the inhomogeneous case, the discrete time case and general diffusion processes are all new, which have not ever been pointed out before.

Furthermore, our results also hold for other Markov processes such as the semi-Markov processes [26] and non-Markovian processes.

2. Transient fluctuation theorems of general stochastic processes

The proof here is enlightened by [20].

Consider a coordinate process $\{X_t : t \geqslant 0\}$ on the canonical trajectory space $\{\Omega, \mathcal{F}, \mathbb{P}\}$, where $\Omega \subseteq \{$A map ω from $\mathbb{R}^+$ or $\mathbb{Z}^+$ to a polish space E that is right continuous having left limit$\}$ and $X_t(\omega) = \omega_t$.

Given $T > 0$, define the time-reversal transformation as follows,

$$r : (\Omega, \mathcal{F}) \to (\Omega, \mathcal{F}), \qquad X_t(r\omega) = \lim_{s \uparrow T-t} X_s(\omega), \qquad \forall t \in [0, T],$$

(in discrete time case, $X_r(r\omega) = X_{T-t}(\omega)$) then r is $\mathcal{F}_{[0,T]}$-measurable and invertible with $r^{-1} = r$, where $\mathcal{F}_{[0,T]}$ is the σ-algebra generated by $\{X_t : 0 \leqslant t \leqslant T\}$. We generally suppose that

$$\omega \in \Omega \quad \Longleftrightarrow \quad r\omega \in \Omega,$$

and by the $\pi - \lambda$ theorem [5, pp 447], we have that $\forall A \in \mathcal{F}_{[0,T]} \Longleftrightarrow rA \in \mathcal{F}_{[0,T]}$, where $rA = \{r\omega : \omega \in A\}$.

Denote $\mathbb{P}_{[0,T]}$ as the distribution of $\{X_t(\omega) : t \in [0, T]\}$, and define $\mathbb{P}^-_{[0,T]} = r\mathbb{P}_{[0,T]}$, which is just the distribution of $\{X_{T-t}(\omega) : t \in [0, T]\}$. So $\forall A \in \mathcal{F}_{[0,T]}$, one has $\mathbb{P}_{[0,T]}(A) = \mathbb{P}^-_{[0,T]}(rA)$.

Here is the key lemma of our proof.

Lemma 2.1. *If* $\mathbb{P}_{[0,T]}$ *and* $\mathbb{P}^-_{[0,T]}$ *are absolutely continuous with respect to each other, then*

$$\frac{\mathrm{d}\mathbb{P}_{[0,T]}}{\mathrm{d}\mathbb{P}^-_{[0,T]}}(\omega) = \frac{\mathrm{d}\mathbb{P}^-_{[0,T]}}{\mathrm{d}\mathbb{P}_{[0,T]}}(r\omega), \; a.e.$$

Proof. $\forall A \in \mathcal{F}_{[0,T]}$, from $\mathbb{P}_{[0,T]}(A) = \mathbb{P}^-_{[0,T]}(rA)$ and the property of the Radon–Nikodym derivative [5, pp 480] follows:

$$\begin{aligned}
\int_A \left[\frac{\mathrm{d}\mathbb{P}^-_{[0,T]}}{\mathrm{d}\mathbb{P}_{[0,T]}}(r\omega)\right] \mathrm{d}\mathbb{P}^-_{[0,T]}(\omega) &= \int_{r^{-1}A} \left[\frac{\mathrm{d}\mathbb{P}^-_{[0,T]}}{\mathrm{d}\mathbb{P}_{[0,T]}}(r(r^{-1}\omega))\right] \mathrm{d}\mathbb{P}_{[0,T]}(\omega) \\
&= \int_{rA} \left[\frac{\mathrm{d}\mathbb{P}^-_{[0,T]}}{\mathrm{d}\mathbb{P}_{[0,T]}}(\omega)\right] \mathrm{d}\mathbb{P}_{[0,T]}(\omega) \\
&= \int_{rA} \mathrm{d}\mathbb{P}^-_{[0,T]}(\omega) \\
&= \mathbb{P}^-_{[0,T]}(rA) \\
&= \mathbb{P}_{[0,T]}(A).
\end{aligned}$$

Finally, from the definition of the Radon–Nikodym derivative [5, pp 220] follows the desired result. □

Define $W_T = \log \frac{\mathrm{d}\mathbb{P}_{[0,T]}}{\mathrm{d}\mathbb{P}^-_{[0,T]}}(\omega)$, and regard $\frac{W}{T}$ as the sample entropy production. It is important to emphasize that W_T is equal to the 'action functional' defined by Lebowitz and Spohn [23]) up to boundary terms of the initial distribution of X_0 and the final distribution of X_T. In fact it is the presence of these boundary terms to ensure the validity of the transient fluctuation theorem. A recent work investigating this problem in homogeneous Markov processes, which is of physical relevance, suggests that these boundary terms in fact cannot be neglected [31].

Theorem 2.2 (transient fluctuation theorem)**.** *Under the condition of lemma 2.1, for each* $z \in \mathbb{R}$,

$$\mathbb{P}_{[0,T]}\left(\frac{W_T}{T} = z\right) = e^{Tz}\mathbb{P}_{[0,T]}\left(\frac{W_T}{T} = -z\right).$$

Proof

$$\begin{aligned}
\mathbb{P}_{[0,T]}\left(\frac{W_T}{T} = z\right) &= \mathbb{P}_{[0,T]}\left(\frac{d\mathbb{P}_{[0,T]}}{d\mathbb{P}^-_{[0,T]}}(\omega) = e^{Tz}\right)\\
&= e^{Tz}\mathbb{P}^-_{[0,T]}\left(\frac{d\mathbb{P}_{[0,T]}}{d\mathbb{P}^-_{[0,T]}}(\omega) = e^{Tz}\right)\\
&= e^{Tz}\mathbb{P}_{[0,T]}\left(\frac{d\mathbb{P}_{[0,T]}}{d\mathbb{P}^-_{[0,T]}}(r\omega) = e^{Tz}\right)\\
&= e^{Tz}\mathbb{P}_{[0,T]}\left(\frac{d\mathbb{P}^-_{[0,T]}}{d\mathbb{P}_{[0,T]}}(\omega) = e^{Tz}\right)\\
&= e^{Tz}\mathbb{P}_{[0,T]}\left(\frac{d\mathbb{P}_{[0,T]}}{d\mathbb{P}^-_{[0,T]}}(\omega) = e^{-Tz}\right)\\
&= e^{Tz}\mathbb{P}_{[0,T]}\left(\frac{W_T}{T} = -z\right).
\end{aligned}$$

□

Remark 2.3. Indeed the derivation of theorem 2.2 is essentially identical to that given in [19, pp 39] and [32]. If W_T takes values in a continuous set, then $\mathbb{P}_{[0,T]}\big(\frac{W_T}{T} = z\big)$ in the previous theorem must be regarded as the probability density rather than the probability itself.

In case one can divide over, the above equality can be written as

$$\frac{\mathbb{P}_{[0,T]}\big(\frac{W_T}{T} = z\big)}{\mathbb{P}_{[0,T]}\big(\frac{W_T}{T} = -z\big)} = e^{Tz}.$$

Such an equality is called the transient fluctuation theorem by Evans *et al* [1, 7, 35–38].

Roughly speaking, the transient fluctuation theorem gives a formula for the probability ratio that the sample entropy production rate $\frac{W_T}{T}$ takes a value z to that of $-z$, and the ratio is e^{Tz}.

3. Applications

3.1. Homogeneous Markov chains

This subsection is rewritten from [20].

3.1.1. The case of discrete time parameter. Let $\xi = \{\xi_n : n \in \mathbb{Z}^+\}$ be a positive recurrent irreducible discrete time Markov chain with denumerable state space S and transition probability matrix $P = (p_{ij})_{i,j\in S}$ on its canonical orbit space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\Omega = S^{\mathbb{Z}}$, $\mathcal{F} = \sigma\{\xi_n : n \in \mathbb{Z}^+\}$ and $\mathbb{P}$ is the distribution of ξ. Its initial distribution $\nu(0) = \{\nu_i(0)\}_{i\in S}$ may be not necessarily its invariant probability distribution $\Pi = \{\pi_i\}_{i\in S}$. Denote the distribution of ξ_n by $\nu(n) = \{\nu_i(n)\}_{i\in S}$.

Suppose that the initial distribution $\nu(0)$ satisfies $\nu_i(0) > 0, \forall i \in S$. For each $n \in \mathbb{Z}^+$ and $k \in \mathbb{N}$, denote by $\mathbb{P}_{[n,n+k]}$ and $\mathbb{P}^-_{[n,n+k]}$ respectively the restrictions of $\mathbb{P}$ and $\mathbb{P}^-$ on $\mathcal{F}_n^{n+k} = \sigma(\xi_m : n \leqslant m \leqslant n+k)$. Assume that the transition matrix P satisfies the condition

$$p_{ij} > 0 \Leftrightarrow p_{ji} > 0, \qquad \forall i, j \in S. \tag{1}$$

The following lemma is analogous to proposition 2.1 in [20].

Lemma 3.1. *Under the condition (1), $\mathbb{P}_{[n,n+k]}$ and $\mathbb{P}^-_{[n,n+k]}$ are absolutely continuous with respect to each other, and the Radon–Nikodym derivative is given by*

$$\frac{\mathrm{d}\mathbb{P}_{[n,n+k]}}{\mathrm{d}\mathbb{P}^-_{[n,n+k]}}(\omega) = \frac{\nu_{\xi_n(\omega)}(n) p_{\xi_n(\omega)\xi_{n+1}(\omega)} \cdots p_{\xi_{n+k-1}(\omega)\xi_{n+k}(\omega)}}{\nu_{\xi_{n+k}(\omega)}(n) p_{\xi_{n+k}(\omega)\xi_{n+k-1}(\omega)} \cdots p_{\xi_{n+1}(\omega)\xi_n(\omega)}}, \qquad \mathbb{P} - \text{a.s.}$$

Thus ξ satisfies the condition of theorem 2.2. Its entropy production rate is measure-theoretically defined as

$$e_p = \lim_{k \to +\infty} \frac{1}{k} E^{\mathbb{P}} \log \frac{\mathrm{d}\mathbb{P}_{[n,n+k]}}{\mathrm{d}\mathbb{P}^-_{[n,n+k]}} = \frac{1}{2} \sum_{i,j \in S} (\pi_i p_{ij} - \pi_j p_{ji}) \log \frac{\pi_i p_{ij}}{\pi_j p_{ji}}.$$

Denote $W_n(\omega) = \log \frac{\mathrm{d}\mathbb{P}_{[0,n]}}{\mathrm{d}\mathbb{P}^-_{[0,n]}}(\omega)$, then $\frac{W_n(\omega)}{n}$ can be regarded as the time-averaged entropy production rate of the sample trajectory ω of the stochastic system modelled by the Markov chain ξ. It holds that

$$\mathbb{P}\left(\frac{W_n}{n} = z\right) = \mathrm{e}^{nz} \mathbb{P}\left(\frac{W_n}{n} = -z\right). \tag{2}$$

Since S is denumerable, $\frac{W_n}{n}$ only takes a denumerable number of values and both sides of the above equality may simultaneously be equal to zero.

If the Markov chain ξ is not reversible, then for $z > 0$ in a certain range, the sample entropy production rate $\frac{W_n}{n}$ has a positive probability to take the value $z > 0$ as well as the value $-z$, but the fluctuation theorem tells that the former probability is greater, which accords with the second law of thermodynamics.

3.1.2. The case of the continuous time parameter. As in the discrete time case, the transient fluctuation theorem holds for stationary or non-stationary Markov chains with a continuous time parameter.

Let $\xi = \{\xi_t : t \in \mathbb{R}^+\}$ be an irreducible Markov chain with finite state space $S = \{1, \ldots, d\}$, and conservative transition density matrix $Q = (q_{ij})_{i,j \in S}$ on its canonical orbit space $(\Omega, \mathcal{F}, \mathbb{P})$ consisting of trajectories that are right continuous with left limits. Its initial distribution $\nu(0) = \{\nu_i(0)\}_{i \in S}$ may be not necessarily its invariant probability distribution $\Pi = \{\pi_i\}_{i \in S}$. We assume that the initial distribution $\nu(0)$ satisfies $\nu_i(0) > 0, \forall i \in S$, and denote the distribution of ξ_t by $\nu(t) = \{\nu_i(t)\}_{i \in S}$.

Similarly, as in the discrete time case, assume that the transition density matrix Q satisfies the condition

$$q_{ij} > 0 \Leftrightarrow q_{ji} > 0, \qquad \forall i, \; j \in S. \tag{3}$$

For each $s \in \mathbb{R}^+$ and $t \in \mathbb{R}^+$, denote by $\mathbb{P}_{[s,s+t)}$ and $\mathbb{P}^-_{[s,s+t)}$ respectively the restrictions of $\mathbb{P}$ and $\mathbb{P}^-$ on $\mathcal{F}_s^{s+t} = \sigma(\xi_u : s \leqslant u < s+t)$.

The following lemma is analogous to proposition 3.1 in [20].

Lemma 3.2. *Under the condition (3), $\mathbb{P}_{[s,s+t)}$ and $\mathbb{P}^-_{[s,s+t)}$ are absolutely continuous with respect to each other, and the Radon–Nikodym derivative is given by*

$$\frac{\mathrm{d}\mathbb{P}_{[s,s+t)}}{\mathrm{d}\mathbb{P}^-_{[s,s+t)}}|_{A_{i_0 i_1\cdots i_n}(t)} = \frac{\nu_{i_0}(s)q_{i_0 i_1}\cdots q_{i_{n-1}i_n}}{\nu_{i_n}(s)q_{i_n i_{n-1}}\cdots q_{i_1 i_0}}, \qquad \mathbb{P}\text{−a.s.},$$

where $A_{i_0 i_1\cdots i_n}(t) = \{\omega \in \Omega: \omega$ jumps n times in $[s, s+t)$, and the states are $i_0, \ldots, i_n$ in turn$\}$.

Thus ξ satisfies the condition of theorem 2.2. The entropy production rate e_p of the Markov chain ξ can be measure-theoretically defined by

$$e_p \overset{\text{def}}{=} \lim_{t\to+\infty} \frac{1}{t} H(\mathbb{P}_{[0,t)}, \mathbb{P}^-_{[0,t)}) = \lim_{t\to+\infty} \frac{1}{t} E^{\mathbb{P}} \log \frac{\mathrm{d}\mathbb{P}_{[0,t)}}{\mathrm{d}\mathbb{P}^-_{[0,t)}} = \frac{1}{2}\sum_{i,j\in S} (\pi_i q_{ij} - \pi_j q_{ji}) \log \frac{\pi_i q_{ij}}{\pi_j q_{ji}},$$

or when $\nu(0) = \Pi$, equivalently, as in [29, 30], by

$$e_p \overset{\text{def}}{=} \lim_{t\downarrow 0+} \frac{1}{t} H(\mathbb{P}_{[s,s+t)}, \mathbb{P}^-_{[s,s+t)}) = \lim_{t\downarrow 0+} \frac{1}{t} E^{\mathbb{P}} \log \frac{\mathrm{d}\mathbb{P}_{[s,s+t)}}{\mathrm{d}\mathbb{P}^-_{[s,s+t)}},$$

where $s \in \mathbb{R}^+$ is arbitrarily fixed. The equivalence is a direct corollary of theorem 10.4 in Varadhan [39].

Denote $W_t = \log \frac{\mathrm{d}\mathbb{P}_{[0,t)}}{\mathrm{d}\mathbb{P}^-_{[0,t)}}$, then $\frac{W_t(\omega)}{t}$ can be regarded as the time-averaged entropy production rate of the sample trajectory ω of the stochastic system modelled by the Markov chain ξ. It holds that

$$\mathbb{P}\left(\frac{W_t}{t} = z\right) = \mathrm{e}^{tz}\mathbb{P}\left(\frac{W_t}{t} = -z\right), \qquad \forall t > 0, z \in \mathbb{R}.$$

3.2. Inhomogeneous Markov chains

3.2.1. Time-periodic inhomogeneous Markov chain. This subsection is recapitulated from [11]. In that paper, we consider periodically inhomogeneous Markov chains, which can be regarded as a simple version of the physical model—Brownian motors [33].

Let $\xi = \{\xi_n : n = 0, 1, 2, \ldots\}$ be a *periodically inhomogeneous Markov chain* on its canonical orbit space $\left(\Omega = \prod_{\mathbb{Z}^+} S, \mathcal{F}, \mathbb{P}\right)$ with denumerable state space S and transition probability matrix $P(m, m+1) = (p_{ij}(m, m+1))_{i,j\in S}$ (we also write it as $P^m = \left(p^m_{ij}\right)_{i,j\in S}$ instead for simplicity), where $p_{ij}(m,n) = \mathbb{P}(\xi_n = j|\xi_m = i), \forall m \leqslant n$. There exists a positive integer T such that

$$p_{ij}(m, m+1) = p_{ij}(m+T, m+T+1), \qquad \forall i, j \in S, \qquad m \in \mathbb{Z}^+. \tag{4}$$

Denote the distribution of ξ_n by $\nu(n) = \{\nu_i(n)\}_{i\in S}$.

For each fixed $k = 0, 1, \ldots, T-1$, the *entropy production rate* e^k_p *of index* k of the periodically inhomogeneous Markov chain $\xi = \{\xi_n : n = 0, 1, 2, \ldots\}$ is measure-theoretically defined as

$$\begin{aligned} e^k_p &= \lim_{n\to\infty} \frac{1}{nT} H(\mathbb{P}_{[k,nT+k]}, \mathbb{P}^-_{[k,nT+k]}) \\ &= \lim_{n\to\infty} \frac{1}{nT} E^{\mathbb{P}} \log \frac{\mathrm{d}\mathbb{P}_{[k,nT+k]}}{\mathrm{d}\mathbb{P}^-_{[k,nT+k]}}(\omega) \\ &= \frac{1}{2T} \sum_{i_0,i_1,\cdots,i_T\in S} \left\{ \left[\pi^k_{i_0} p^k_{i_0 i_1} p^{k+1}_{i_1 i_2} \cdots p^{k+T-1}_{i_{T-1} i_T} - \pi^k_{i_T} p^k_{i_T i_{T-1}} p^{k+1}_{i_{T-1} i_{T-2}} \cdots p^{k+T-1}_{i_1 i_0}\right] \right. \\ &\quad \left. \times \log \frac{\pi^k_{i_0} p^k_{i_0 i_1} p^{k+1}_{i_1 i_2} \cdots p^{k+T-1}_{i_{T-1} i_T}}{\pi^k_{i_T} p^k_{i_T i_{T-1}} p^{k+1}_{i_{T-1} i_{T-2}} \cdots p^{k+T-1}_{i_1 i_0}} \right\}, \end{aligned} \tag{5}$$

where $\mathbb{P}_{[k,nT+k]}$ is the distribution of $(\xi_k, \xi_{k+1}, \ldots, \xi_{nT+k})$, $\mathbb{P}^-_{[k,nT+k]}$ is the distribution of $(\xi_{nT+k}, \xi_{nT+k-1}, \ldots, \xi_{k+1}, \xi_k)$ and $\Pi = \{\pi^0, \pi^1, \ldots, \pi^{T-1}\}$ is the periodically stationary distribution of ξ.

The following lemma is analogous to lemma 3.5 in [11].

Lemma 3.3. *Suppose that* $\forall i_0, i_1, \ldots, i_T \in S$, $p^k_{i_0 i_1} p^{k+1}_{i_1 i_2} \cdots p^{k+T-1}_{i_{T-1} i_T} > 0$ *if and only if* $p^k_{i_T i_{T-1}} p^{k+1}_{i_{T-1} i_{T-2}} \cdots p^{k+T-1}_{i_1 i_0} > 0$. *For arbitrarily fixed* $k \geqslant 0$, *assume that* $\nu_i(k) > 0$ *for all* $i \in S$, *then* $\mathbb{P}_{[k,nT+k]}$ *and* $\mathbb{P}^-_{[k,nT+k]}$ *are absolutely continuous with respect to each other, and* $\mathbb{P}$*-almost everywhere, the Radon–Nikodym derivative is*

$$\frac{\mathrm{d}\mathbb{P}_{[k,nT+k]}}{\mathrm{d}\mathbb{P}^-_{[k,nT+k]}}(\omega) = \frac{\nu_{\xi_k(\omega)}(k) p^k_{\xi_k(\omega)\xi_{k+1}(\omega)} \cdots p^{nT+k-1}_{\xi_{nT+k-1}(\omega)\xi_{nT+k}(\omega)}}{\nu_{\xi_{nT+k}(\omega)}(k) p^k_{\xi_{nT+k}(\omega)\xi_{nT+k-1}(\omega)} \cdots p^{nT+k-1}_{\xi_{k+1}(\omega)\xi_k(\omega)}}. \tag{6}$$

Let $W^k_n(\omega) = \log \frac{\mathrm{d}\mathbb{P}_{[k,nT+k]}}{\mathrm{d}\mathbb{P}^-_{[k,nT+k]}}(\omega)$, then holds the transient fluctuation theorem

$$\mathbb{P}\left(\frac{W^k_n}{nT} = z\right) = \mathrm{e}^{nTz}\mathbb{P}\left(\frac{W^k_n}{nT} = -z\right),$$

which is a generalization of corollary 4.5 in [11].

3.2.2. Discrete-time inhomogeneous Markov chains. The instantaneous reversibility and entropy production of inhomogeneous Markov chains are defined and discussed in [12]. But the transient fluctuation theorem has not been derived yet for this case.

Let $\xi = \{\xi_n : n = 0, 1, 2, \ldots\}$ be an inhomogeneous Markov chain with denumerable state space S and transition probability $p^k_{ij} = \mathbb{P}(\xi_{k+1} = j | \xi_k = i)$. Denote the distribution of ξ_n by $\nu(n) = \{\nu_i(n)\}_{i \in S}$. The instantaneous entropy production e^k_p of the inhomogeneous Markov chain ξ at time k can be measure-theoretically expressed as

$$e^k_p = H(\mathbb{P}_{[k,k+1]}, \mathbb{P}^-_{[k,k+1]}) = \frac{1}{2}\sum_{i,j \in S}\left(\nu_i(k)p^k_{ij} - \nu_j(k)p^k_{ji}\right)\log\frac{\nu_i(k)p^k_{ij}}{\nu_j(k)p^k_{ji}},$$

where $\mathbb{P}_{[k,k+1]}$ is the distribution of (ξ_k, ξ_{k+1}), $\mathbb{P}^-_{[k,k+1]}$ is the distribution of (ξ_{k+1}, ξ_k) and $H(\mathbb{P}_{[k,k+1]}, \mathbb{P}^-_{[k,k+1]})$ is the relative entropy of $\mathbb{P}_{[k,k+1]}$ with respect to $\mathbb{P}^-_{[k,k+1]}$.

Lemma 3.4. *For arbitrarily fixed* $n \geqslant 0$, *suppose that* $\nu_i(n) > 0$, $\forall i \in S$. *Under the condition*

$$p^m_{ij} > 0 \quad \textit{for some} \quad m \in \mathbb{N} \quad \Longleftrightarrow \quad p^m_{ji} > 0 \quad \textit{for all} \quad m \in \mathbb{N},$$

$\mathbb{P}_{[n,n+k]}$ *and* $\mathbb{P}^-_{[n,n+k]}$ *are absolutely continuous with respect to each other, and the Radon–Nikodym derivative is given by*

$$\frac{\mathrm{d}\mathbb{P}_{[n,n+k]}}{\mathrm{d}\mathbb{P}^-_{[n,n+k]}}(\omega) = \frac{\nu_{\xi_n(\omega)}(n) p^n_{\xi_n(\omega)\xi_{n+1}(\omega)} \cdots p^{n+k-1}_{\xi_{n+k-1}(\omega)\xi_{n+k}(\omega)}}{\nu_{\xi_{n+k}(\omega)}(n) p^n_{\xi_{n+k}(\omega)\xi_{n+k-1}(\omega)} \cdots p^{n+k-1}_{\xi_{n+1}(\omega)\xi_n(\omega)}}, \qquad \mathbb{P}-\text{a.s.}$$

Denote $W_n(\omega) = \log \frac{\mathrm{d}\mathbb{P}_{[0,n]}}{\mathrm{d}\mathbb{P}^-_{[0,n]}}(\omega)$, then it also holds that

$$\mathbb{P}\left(\frac{W_n}{n} = z\right) = \mathrm{e}^{nz}\mathbb{P}\left(\frac{W_n}{n} = -z\right). \tag{7}$$

3.2.3. Continuous-time inhomogeneous Markov chains. The instantaneous reversibility and entropy production of continuous-time inhomogeneous Markov chains are defined and discussed in [12]. Such chains have been recently applied to derive Jarzynski's equality [13]. But the transient fluctuation theorem has not been derived yet for this case.

Let $\xi = \{\xi(t): t \geqslant 0\}$ be an inhomogeneous Markov chain with a denumerable state space S and a conservative continuous transition density function $Q(\cdot)$. Denote the distribution of $\xi(t)$ by $\nu(t) = \{\nu_i(t)\}_{i\in S}$.

As in the case of homogeneous Markov chains [34], the *instantaneous entropy production rate* $e_p(s)$ of the inhomogeneous Markov chain $\{\xi(t) : t \geqslant 0\}$ can be measure-theoretically expressed as

$$e_p(s) = \lim_{t\downarrow s} \frac{1}{t-s} H(\mathbb{P}_{[s,t]}, \mathbb{P}^-_{[s,t]}) = \frac{1}{2}\sum_{i,j\in S}[\nu_i(s)q_{ij}(s) - \nu_j(s)q_{ji}(s)] \log \frac{\nu_i(s)q_{ij}(s)}{\nu_j(s)q_{ji}(s)},$$

where $\mathbb{P}_{[s,t]}$ is the distribution of $\{\xi(u): s \leqslant u \leqslant t\}$, $\mathbb{P}^-_{[s,t]}$ is the distribution of $\{\xi(s+t-u): s \leqslant u \leqslant t\}$, and $H(\mathbb{P}_{[s,t]}, \mathbb{P}^-_{[s,t]})$ is the relative entropy of $\mathbb{P}_{[s,t]}$ with respect to $\mathbb{P}^-_{[s,t]}$.

Denote by n_t the number of times that ξ jumps in $[s, t]$. Let $T_0 = s$, $T_1 = \inf\{t > s: \xi(t) \neq \xi(s)\}$, $T_k = \inf\{t > T_{k-1}: \xi(t) \neq \xi(T_{k-1})\}$, and $T_{n_t+1} = t$, then $\forall i_0, i_1, \ldots, i_n \in S$ satisfying $i_k \neq i_{k+1}$, $0 \leqslant k < n$, we can define

$$A_{i_0 i_1\cdots i_n}(t) = \{\omega \in \Omega: n_t(\omega) = n, \xi(s) = i_0, \xi(T_k(\omega)) = i_k, k = 1, \ldots, n\}.$$

The following is lemma 4.2 in [12].

Lemma 3.5. *For arbitrarily fixed* $s \geqslant 0$, *suppose that* $\nu_i(s) > 0$ *for all* $i \in S$. *If the transition density function* $Q(\cdot)$ *satisfies the condition*

$$q_{ij}(u) > 0 \qquad \text{for some} \quad u > 0 \quad \Longleftrightarrow \quad q_{ji}(u) > 0, \quad \forall u > 0,$$

then $\mathbb{P}_{[s,t]}$ *and* $\mathbb{P}^-_{[s,t]}$ *are absolutely continuous with respect to each other. The Radon–Nikodym derivative is a.e.*

$$\left.\frac{\mathrm{d}\mathbb{P}_{[s,t]}}{\mathrm{d}\mathbb{P}^-_{[s,t]}}\right|_{A_{i_0 i_1\cdots i_n}(t)} = \frac{\nu_{i_0}(s)\prod_{k=0}^{n-1} q_{i_k i_{k+1}}(T_{k+1}) \exp\left[\sum_{k=0}^{n}\int_{T_k}^{T_{k+1}} q_{i_k i_k}(u)\,\mathrm{d}u\right]}{\nu_{i_n}(s)\prod_{k=0}^{n-1} q_{i_{n-k} i_{n-k+1}}(s+t-T_{n-k}) \exp\left[\sum_{k=0}^{n}\int_{s+t-T_{k+1}}^{s+t-T_k} q_{i_k i_k}(u)\,\mathrm{d}u\right]}.$$

Theorem 3.6 (Transient fluctuation theorem for inhomogeneous Markov chains)**.** *Let* $W_t = \log \frac{d\mathbb{P}_{[0,t)}}{d\mathbb{P}^-_{[0,t)}}$, *then*

$$\mathbb{P}\left(\frac{W_t}{t} = z\right) = \mathrm{e}^{tz}\mathbb{P}\left(\frac{W_t}{t} = -z\right), \qquad \forall t > 0, z \in \mathbb{R}.$$

3.3. General stationary diffusion processes

The instantaneous reversibility and entropy production of diffusion processes are defined in [19, chapters 3, 4], but the transient fluctuation theorem has not been derived yet for this case.

The non-degenerate diffusion process $\xi = \{\xi_t\}_{t\geqslant 0}$ constructed in [27] and [19, chapter 3] can also be understood as the solution of the following stochastic differential equation:

$$\mathrm{d}\xi_t = \Gamma(\xi_t)\,\mathrm{d}B_t + \bar{b}(\xi_t)\,\mathrm{d}t, \tag{8}$$

where $\Gamma(\cdot)$ is a $d \times m$ matrix, $\bar{b}(\cdot)$ is a vector field on $\mathbb{R}^d$ and $\{B_t\}_{t\geqslant 0}$ is an m-dimensional Brownian motion. Recall that the infinitesimal generator of ξ is

$$\mathcal{A} = \nabla\cdot\left(\frac{1}{2}A(x)\nabla\right) + b(x)\cdot\nabla = \frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial}{\partial x^i}a^{ij}(x)\frac{\partial}{\partial x^j} + \sum_{i=1}^{d} b^i(x)\frac{\partial}{\partial x^i}$$

with locally elliptic $A(x) = (a^{ij}(x)) = \Gamma(x)\Gamma^T(x)$ and $b(x) = (b^i(x)) = \bar{b}(x) - c(x)$, where $c(x) = \frac{1}{2}\nabla A(x)$, namely $c^i(x) = \frac{1}{2}\sum_{j=1}^{d}\frac{\partial a^{ij}}{\partial x^j}$. For simplicity and without loss of generality, it can be thought that $m = d$ and $\Gamma = A^{\frac{1}{2}}$.

The *entropy production rate* e_p of a stationary diffusion process ξ can be measure-theoretically expressed as

$$
\begin{aligned}
e_p &\overset{\text{def}}{=} \lim_{t\to+\infty}\frac{1}{t}H(\mathbb{P}_{[0,t]},\mathbb{P}^-_{[0,t]})\\
&= \frac{1}{2}\int_{\mathbb{R}^d}[2b(x) - A\nabla\log\rho(x)]^T A^{-1}[2b(x) - A\nabla\log\rho(x)]\rho(x)\,\mathrm{d}x,
\end{aligned}
\tag{9}
$$

where $\rho(x)$ is the invariant probability density of the process and $H(\mathbb{P}_{[0,t]},\mathbb{P}^-_{[0,t]})$ is the relative entropy of $\mathbb{P}_{[0,t]}$ with respect to $\mathbb{P}^-_{[0,t]}$.

The following is proposition 4.1.6 in [19], which holds for the general non-degenerate diffusion process [19, subsection 4.1.2].

Lemma 3.7. *For each $t > 0$, the two probability measures $\mathbb{P}_{[0,t]}$ and $\mathbb{P}^-_{[0,t]}$, where $\mathbb{P}_{[0,t]}$ is the distribution of $\{\xi_s : 0 \leqslant s \leqslant t\}$ and $\mathbb{P}^-_{[0,t]}$ is the distribution of $\{\xi_{t-s} : 0 \leqslant s \leqslant t\}$, are equivalent to each other. Moreover, the positive measurable function $\frac{\mathrm{d}\mathbb{P}_{[0,t]}}{\mathrm{d}\mathbb{P}^-_{[0,t]}}$ on $\left(C([0,\infty),\mathbb{R}^d),\mathcal{B}_0^t\right)$ satisfies that for $\mathbb{P}$-almost every $\omega\in\Omega$,*

$$
\begin{aligned}
\frac{\mathrm{d}\mathbb{P}_{[0,t]}}{\mathrm{d}\mathbb{P}^-_{[0,t]}}(\xi.(\omega)) = \exp\Bigg[&\int_0^t (2A^{-1}b - \nabla\log\rho)^T(\xi_s)\,\mathrm{d}\bar{\xi}_s(\omega)\\
&+ \frac{1}{2}\int_0^t (2A^{-1}b - \nabla\log\rho)^T A(2A^{-1}b - \nabla\log\rho)(\xi_s(\omega))\,\mathrm{d}s\Bigg],
\end{aligned}
$$

where $\mathrm{d}\bar{\xi}_s = \mathrm{d}\xi_s - \bar{b}(\xi_s)\,\mathrm{d}s$ and ρ is the invariant probability density of ξ under $\mathbb{P}$.

Theorem 3.8 (transient fluctuation theorem for general diffusion processes)**.** *Let $W_t = \log\frac{\mathrm{d}\mathbb{P}_{[0,t]}}{\mathrm{d}\mathbb{P}^-_{[0,t]}}(\xi.(\omega))$, then for each $z\in\mathbb{R}$,*

$$
\mathbb{P}_{[0,t]}\left(\frac{W_t}{t} = z\right) = \mathrm{e}^{tz}\mathbb{P}_{[0,t]}\left(\frac{W_t}{t} = -z\right).
$$

It is important to note that W_t takes values in a continuous set, so $\mathbb{P}_{[0,t]}\big(\frac{W_t}{t} = z\big)$ in the previous theorem must be regarded as the probability density rather than the probability itself.

Now we will give a brief discussion of how Kurchan's result concerning the fluctuations of external work done on the system can be derived from our result in this subsection.

The following result is a direct corollary of theorem 3.8, which is equivalent to the 'first version of the fluctuation theorem' in Kurchan's work [22], and one can find another approach of proof in [19, pp 65].

Corollary 3.9

$$
E\,\mathrm{e}^{\lambda W_t} = E\,\mathrm{e}^{-(\lambda+1)W_t}, \qquad \forall\lambda\in\mathbb{R}. \tag{10}
$$

Proof

$$
\begin{aligned}
E\mathrm{e}^{\lambda W_t} &= \int \mathrm{e}^{\lambda z}\mathbb{P}_{[0,t]}(W_t = z)\,\mathrm{d}z\\
&= \int \mathrm{e}^{\lambda z}e^{z}\mathbb{P}_{[0,t]}(W_t = -z)\,\mathrm{d}z
\end{aligned}
$$

$$= \int e^{-\lambda z} e^{-z} \mathbb{P}_{[0,t]}(W_t = z)\, dz$$
$$= E\, e^{-(\lambda+1)W_t}. \qquad \square$$

However, from the mathematical point of view, the expression and derivation of the steady-state fluctuation theorem in [22] are not rigorous and clear. That is why Lebowitz and Spohn [23] introduced the language of large deviation theory to express the steady-state fluctuation theorem, although their proof in the case of diffusion processes are not mathematically rigorous either.

What has been rigorously proved is the ergodic theorem for sample entropy production

$$\lim_{t\to\infty} \frac{W_t}{t} = e_p,$$

according to [19, proposition 4.1.8].

For general diffusion processes, let $c_t(\lambda) = \frac{1}{t} \log E e^{\lambda W_t}$. According to the well-known Ellis–Gartner theorem [19, theorem 1.5.2] for a large deviation property, if

(a) each function $c_t(\lambda)$ is finite for all $\lambda \in \mathbb{R}$,

(b) $c(\lambda) = \lim_{t\to\infty} c_t(\lambda)$ exists for all $\lambda \in \mathbb{R}$ and is finite, which is always called the free energy function, and

(c) $c(\lambda)$ is differentiable for all λ, then $\left\{\frac{W_t}{t}\right\}$ has a large deviation property with the rate function (also called entropy function) $I(z) = \sup_{\lambda\in\mathbb{R}}\{\lambda z - c(\lambda)\}$.

Under the above conditions, according to corollary 3.9, $c(\lambda) = c(-1-\lambda)$, which yields $I(z) = I(-z) - z$. This is just the steady-state fluctuation theorem of the Lebowitz–Spohn type, which is the mathematical counterpart of Kurchan's work.

But it is very difficult to prove these conditions (a), (b) and (c), although they are natural for statistical mechanical applications.

References

[1] Ayton G and Evans D J 1999 On the asymptotic convergence of the transient and steady-state fluctuation theorems *J. Stat. Phys.* **97** 811–5

[2] Crooks G E 1998 Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems *J. Stat. Phys.* **90** 1481–7

[3] Crooks G E 1999 Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences *Phys. Rev.* E **60** 2721–6

[4] Crooks G E 2000 Path-ensemble averages in systems driven far from equilibrium *Phys. Rev.* E **61** 2361–6

[5] Durrett R 1996 *Probability: Theory and Examples* 2nd edn (Pacific Grove, CA: Duxbury Press)

[6] Evans D J, Cohen E G D and Morriss G P 1993 Probability of second law violation in steady flows *Phys. Rev. Lett.* **71** 2401–4

[7] Evans D J and Searles D J 1994 Equilibrium microstates which generate second law violating steady states *Phys. Rev.* E **50** 1645–8

[8] Gallavotti G 1995 Reversible Anosov diffeomorphisms and large deviations *Math. Phys. Electron. J.* **1** 1–12

[9] Gallavotti G 1996 Chaotic hypothesis: Onsager reciprocity and fluctuation-dissipation theorem *J. Stat. Phys.* **84** 899–926

[10] Gallavotti G and Cohen E G D 1995 Dynamical ensembles in stationary states *J. Stat. Phys.* **80** 931–70

[11] Ge H, Jiang D Q and Qian M 2006 A simple discrete model of Brownian motors: time-periodic Markov chains *J. Stat. Phys.* **123** 831–59

[12] Ge H, Jiang D Q and Qian M 2006 Reversibility and entropy production of inhomogeneous Markov chains *J. Appl. Prob.* **43** 1028–43

[13] Ge H and Qian M 2007 Generalized Jarzynski's equality in inhomogeneous Markov chains *J. Math. Phys.* **48** 053302

[14] Hasegawa H 1976 *Prog. Theor. Phys.* **55** 90
Hasegawa H 1976 *Prog. Theor. Phys.* **56** 44
Hasegawa H 1977 *Prog. Theor. Phys.* **57** 1523
Hasegawa H 1977 *Prog. Theor. Phys.* **58** 128

[15] Jarzynski C 1997 Nonequilibrium equality for free energy differences *Phys. Rev. Lett.* **78** 2690–3
[16] Jarzynski C 1997 Equilibrium free-energy differences from nonequilibrium measurements: a master-equation approach *Phys. Rev.* E **56** 5018–35
[17] Jarzynski C 1999 Microscopic analysis of Clausius-duhem processes *J. Stat. Phys.* **96** 415–27
[18] Jarzynski C 2000 Hamiltonian derivation of a detailed fluctuation theorem *J. Stat. Phys.* **98** 77–102
[19] Jiang D Q, Qian M and Qian M P 2004 *Mathematical Theory of Nonequilibrium Steady States: On the Frontier of Probability and Dynamical Systems (Lect. Notes Math.* vol 1833*)* (Berlin: Springer)
[20] Jiang D Q, Qian M and Zhang F X 2003 Entropy production fluctuations of finite Markov chains *J. Math. Phys.* **44** 4176–88
[21] Keizer J 1987 *Statistical Thermodynamics of Nonequilibrium Processes* (New York: Springer)
[22] Kurchan J 1998 Fluctuation theorem for stochastic dynamics *J. Phys. A: Math. Gen.* **31** 3719–29
[23] Lebowitz J L and Spohn H 1999 A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics *J. Stat. Phys.* **95** 333–65
[24] Maes C 1999 The fluctuation theorem as a Gibbs property *J. Stat. Phys.* **95** 367–92
[25] Nicolis G and Prigogine I 1977 *Self-organization in Nonequilibrium Systems* (New York: Academic)
[26] Qian H and Wang H 2007 On detailed balance and reversibility of semi-Markov processes and single-molecule enzyme kinetics *J. Math. Phys.* **48** 013303
[27] Qian H, Qian M and Tang X 2002 Thermodynamics of the general diffusion process: time-reversibility and entropy production *J. Stat. Phys.* **107** 1129–41
[28] Qian M P and Qian M 1982 Circulation for recurrent Markov chains *Z. Wahrscheinlichkeitstheor. Verwandte Geb.* **59** 203–10
[29] Qian M P and Qian M 1987 The entropy production and reversibility of Markov processes *Probability Theory and Applications, Proc. 1st World Congress of the Bernoulli Society, Tashkent, USSR 1986* vol 1 ed A Prohorov Yu and V V Sazonov (Utrecht, The Netherlands: VNU Science) pp 307–16
[30] Qian M P, Qian M and Gong G L 1991 The reversibility and the entropy producton of Markov processes *Contemp. Math.* **118** 255–61
[31] Puglisi A, Rondoni L and Vulpiani A 2006 Relevance of initial and final conditions for the fluctuation relation in Markov processes *J. Stat. Mech.* P08010
[32] Reid J C, Sevick E M and Evans D J 2005 A unified description of two theorems in non-equilibrium statistical mechanics: the fluctuation theorem and the work relation *Europhys. Lett.* **72** 726–32
[33] Reimann P 2002 Brownian motors: noisy transport far from equilibrium *Phys. Rep.* **361** 57–265
[34] Schnakenberg J 1976 Network theory of microscopic and macroscopic behaviour of master equation systems *Rev. Mod. Phys.* **48** 571–85
[35] Searles D J and Evans D J 1999 Fluctuation theorem for stochastic systems *Phys. Rev.* E **60** 159–64
[36] Searles D J and Evans D J 2000 The fluctuation theorem and Green–Kubo relations *J. Chem. Phys.* **112** 9727–35
[37] Searles D J and Evans D J 2000 Ensemble dependence of the transient fluctuation theorem *J. Chem. Phys.* **113** 3503–9
[38] Searles D J and Evans D J 2001 Fluctuation theorem for heat flow *Int. J. Thermophys.* **22** 123–34
[39] Varadhan S R S 1984 *Large Deviations and Applications* (Philadelphia, PA: Society for Industrial and Applied Mathematics)