

UCLA

UCLA Previously Published Works

Title

Fluctuation Theorems for Entropy Production and Heat Dissipation in Periodically Driven Markov Chains

Permalink

<https://escholarship.org/uc/item/7z75481s>

Journal

Journal of Statistical Physics: 1, 137(1)

ISSN

1572-9613

Authors

Shargel, Benjamin Hertz
Chou, Tom

Publication Date

2009-10-01

DOI

10.1007/s10955-009-9836-8

Peer reviewed

Fluctuation Theorems for Entropy Production and Heat Dissipation in Periodically Driven Markov Chains

Benjamin Hertz Shargel · Tom Chou

Received: 29 March 2009 / Accepted: 17 September 2009 / Published online: 26 September 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract Asymptotic fluctuation theorems are statements of a Gallavotti-Cohen symmetry in the rate function of either the time-averaged entropy production or heat dissipation of a process. Such theorems have been proved for various general classes of continuous-time deterministic and stochastic processes, but always under the assumption that the forces driving the system are time independent, and often relying on the existence of a limiting ergodic distribution. In this paper we extend the asymptotic fluctuation theorem for the first time to inhomogeneous continuous-time processes without a stationary distribution, considering specifically a finite state Markov chain driven by periodic transition rates. We find that for both entropy production and heat dissipation, the usual Gallavotti-Cohen symmetry of the rate function is generalized to an analogous relation between the rate functions of the original process and its corresponding backward process, in which the trajectory and the driving protocol have been time-reversed. The effect is that spontaneous positive fluctuations in the long time average of each quantity in the forward process are exponentially more likely than spontaneous *negative* fluctuations in the *backward* process, and vice-versa, revealing that the distributions of fluctuations in universes in which time moves forward and backward are related. As an additional result, the asymptotic time-averaged entropy production is obtained as the integral of a periodic entropy production rate that generalizes the constant rate pertaining to homogeneous dynamics.

Keywords Fluctuation theorem · Large deviations · Entropy production

1 Introduction

Following the initial computer simulations of Evans et al. [9] and the pioneering paper by Gallavotti and Cohen [12], the study of fluctuation theorems has led to a fascinating conflu-

B.H. Shargel (✉)
Department of Mathematics, UCLA, Los Angeles, CA 90095-1766, USA
e-mail: shargel@math.ucla.edu

T. Chou
Departments of Mathematics and Biomathematics, UCLA, Los Angeles, CA 90095-1766, USA

ence of irreversible thermodynamics, stochastic processes and large deviation theory. Unlike its original role in statistical physics of formalizing the thermodynamic limit for equilibrium ensembles, the contribution large deviation theory makes in this context is to characterize the fluctuations in the long time average of the entropy production or heat dissipation of a stochastic process, which models a physical system whose number of degrees of freedom or incomplete description makes a deterministic treatment infeasible. To see how this characterization comes about and these disparate fields fit together, consider an stochastic process ξ_t over a state space S with law $\mu(x, t)$, which for simplicity we take to be either a density or discrete distribution. The Gibbs entropy of the process, viewed as an ensemble of paths (or a measure over that ensemble), is $-\int_S \log \mu(x, t) \mu(dx, t)$, which leads one to identify the entropy along a single stochastic trajectory as $-\log \mu(\xi_t, t)$ [33]. The time derivative of this quantity equals the difference between the rate of entropy produced by the stochastic particle and the rate of entropy, or heat (divided by a nonphysical temperature), dissipated to its environment. Maes [25] as well as Lebowitz and Spohn [24] further recognized that under general circumstances the total entropy production equals the logarithmic Radon-Nikodym derivative of the forward path measure P governing the process with respect to its corresponding backward path measure P^B which, under Crooks' more general definition [4], is obtained by time-reversing all temporal inhomogeneities driving the process and composing with a path-reversal transformation. As $P^B(\omega)$ equals the probability of observing a path ω unfold *in reverse* as time runs *backward* from time t to 0, the entropy production may be interpreted as the log likelihood of observing ω in a universe in which time runs forward as opposed to backward. The existence of the derivative (i.e. the equivalence of P and P^B) depends on a condition called at times dynamic reversibility [27] or ergodic consistency [8], which ensures that the time-reversal of any trajectory realizable in the forward process is realizable in the backward process.

The Radon-Nikodym definition for entropy production was later justified thoroughly by Maes and collaborators for a wide range of deterministic and stochastic processes [26, 27]. It is nearly identical to that of the dissipation function Ω in deterministic mechanics [8, 34], except that the forward and backward measures in that case put full mass on the constant energy manifold of trajectories obeying Hamilton's equations. The logarithmic derivative is nonzero for nonstationary processes, which model physical systems evolving far from equilibrium, as well as for stationary ones violating detailed balance, modeling systems in a nonequilibrium steady state. Its expectation under P equals the relative entropy of P with respect to P^B , which is always nonnegative, consistent with a weak reading of the second law of thermodynamics. The time-extensive microscopic heat dissipation Lebowitz and Spohn termed an action functional, a quantity that is equal to the above logarithmic derivative up to the difference of boundary terms $-\log \frac{\mu(\xi_t, t)}{\mu(\xi_0, 0)}$, precisely the net change in system entropy.

Recent work has exploited the thermodynamic framework above by proving that either the time-averaged entropy production or heat dissipation satisfies a large deviation principle, whose corresponding rate function satisfies the same symmetry as the one proved by Gallavotti and Cohen [12] to hold for the time-averaged phase space contraction of chaotic dynamical systems. We refer to this as an *asymptotic fluctuation theorem* (AFT) or, simply, a fluctuation theorem, and distinguish it from transient fluctuation theorems, which hold at finite times but are not large deviation results. Kurchan first proved an AFT [21] for the entropy production of Langevin processes under the assumption of nondegeneracy for the maximum eigenvalue of their evolution operator, and Lebowitz and Spohn then proved it [24] for the heat dissipation of time-homogeneous Markov chains and Itô diffusions, whose assumed limiting stationary distribution guarantees this condition. The cases of continuous

and discrete time Markov chains were handled in a rigorous fashion by Jiang et al. [18, 19], and Ge et al. extended the discrete time case to include time-periodic inhomogeneities in the transition matrix [16]. Maes took a different approach from the others, studying the dynamics of finite volume Gibbs states on \mathbb{Z}^d under general space-time potentials [25].

Interestingly, it has been shown that unlike their entropy production, the heat dissipation of Langevin processes does not satisfy an AFT, at least in the conventional sense [39, 40], displaying importance of the boundary terms distinguishing these quantities. Along similar lines, Rákos and Harris have recently shown [30] that infinite state spaces can result in the divergence of the boundary terms, causing a breakdown of the Gallavotti-Cohen symmetry. It is unclear, however, whether this breakdown, which arises from the failure of the Hamiltonian operator to satisfy Kurchan's nondegeneracy requirement, is really a function of the state space's cardinality as much as its non-compactness (the state space of a Markov chain being effectively endowed with the discrete metric). In the case of deterministic dynamics, Bonetto et al. have argued that apparent violations of AFTs by particle systems with singularities in the interparticle potential (owing, for example, to hard cores) can be corrected by subtracting "unphysical" singular terms from the phase space contraction rate [2].

While so much effort recently has gone into circumscribing the range of applicability of the AFT, the purpose of our present work is to extend it for the first time to the case of inhomogeneous, continuous time dynamics. We consider in particular a Markov chain on a finite state space, whose infinitesimal generator is a continuous and periodic function of time but only required to be irreducible at a single moment. The finiteness of the state space ensures that the complexity of the model is isolated within the time dimension, avoiding in particular the issues raised in [30]. Such a process can be used to model phenomena as diverse as the fluctuation-driven transport of molecular motors [1], stochastic resonance in lasers and neuron firing [13], quasienergy banding in periodically-driven mesoscopic electric circuits [3], and seasonality in population dynamics [31], as well as periodically-driven deterministic processes amenable to coarse-graining. Continuous time models of all of these phenomena were previously outside the scope of AFTs, all of which had been proved under the assumption of homogeneous dynamics, because they rely fundamentally on a time-dependent protocol driving the process. Indeed, few systems in nature operate within a static environment, and so to gain true scientific relevance AFTs must ultimately accommodate time-inhomogeneities. While our assumption of periodicity remains a restriction to potential applications, we believe that having laid out in this paper the mathematical issues involved in introducing time-dependent driving, our arguments can serve as a blueprint for future work that seeks to loosen this restriction.

Whereas the authors mentioned above have proved AFTs for either the action functional or entropy production of a process alone, we prove for both, finding that the absence of boundary terms in the action functional simplifies the derivation of its free energy but complicates that of its fluctuation symmetry. The fluctuation symmetry for both quantities takes the form $I(z) - I^B(-z) = -z$, where I is the rate function under the forward process and I^B under the backward process, which reduces to the usual Gallavotti-Cohen symmetry when the driving protocol is symmetric inside every driving period and, hence, has no temporal orientation. Its interpretation is that spontaneous positive fluctuations in the long time average of each quantity in one process is exponentially more likely than spontaneous *negative* fluctuations in the other, a relation that is symmetric with respect to the two processes.

We also derive the almost sure asymptotic time-averaged entropy production as the integral $\int_0^T e_p(s) ds$, where $e_p(s)$ is a T -periodic instantaneous entropy production rate, with $T \in \mathbb{R}^+$ the period of the driving. This expression generalizes existing ones known for homogeneous [14, 18] and periodically inhomogeneous, but discrete time [16], chains. Our

proofs are guided strategically by those of Lebowitz and Spohn, but, as discussed above, the inhomogeneity of the process necessitates more involved, and rigorous, arguments. The virtue of this is that our proofs, unlike those of Lebowitz and Spohn, do not rely explicitly on the existence of a stationary distribution for the process (indeed, only a periodic quasi-stationary distribution exists), raising the open question of how insensitive the existence of an AFT is to the details of the asymptotic regime of the process. Put another way, how far can our assumption of periodicity on the driving rates be relaxed but still guarantee an AFT? We address this question in the final section of the paper, proving that uniform continuity and boundedness of the rates alone, which are implied by the assumptions described in Sect. 2, are not sufficient. This shows that any minimal set of conditions on the rates are closer to those assumed in this paper than one might initially suspect.

A number of theoretical and experimental results already exist for periodically-driven stochastic processes. Integral and transient fluctuation theorems have been proved for periodically-driven two-state Markov chains [32, 37], quantum systems [22], as well as classical harmonic oscillators modeled by a Langevin process [20, 35]. It must be emphasized that these studies differ from ours because they do not prove large deviations results. As the technicalities we deal with in this paper reveal, the transition from transient to asymptotic fluctuation theorem, even for processes as simple as finite state Markov chains, is not an automatic or obvious one, but depends on the details of the driving protocol.

Similarly, in spite of the close relation between discrete and continuous time Markov chains, our results cannot be obtained from those of Ge et al. The set of transition probabilities in their case is finite, and the uniformity of the waiting times between jumps enables the inhomogeneous chain on S to be re-represented as a homogeneous chain on the larger (but still finite) space $S^{T'}$, where $T' \in \mathbb{N}$ is the period of the transition probabilities. Our set of transition rates is uncountable, on the contrary, and is sampled randomly by the process, ensuring no convenient reduction to a homogeneous problem. It is possible, of course, that with the proper scaling of the transition probabilities (under which the periodicity T' would diverge), the continuous time path measure may be obtained as the weak limit of the discrete time path measures, analogous to Donsker's theorem for Brownian motion. However, as has been mentioned, the rate functions for the heat dissipation and entropy production of the continuous process do not simply inherit the fluctuation symmetry of their discrete approximants. Instead, the internal symmetry of each rate function is replaced by a symmetry between it and its counterpart under the corresponding backward process.

The rest of the paper is organized as follows. We begin by making the appropriate definitions, and then derive a backward equation for the moment generating function of the action functional, whose fundamental solution we obtain using both Floquet theory and the time-ordered exponential operator. This enables us to identify the free energy of the action functional as the nondegenerate principal Floquet eigenvalue of the fundamental solution, thereby proving the existence and strict convexity of its Legendre-Fenchel transform, the rate function. We then show that the free energies of the action functional under the forward and backward path measures satisfy a symmetry relation analogous to the one found by Lebowitz and Spohn in the case of time-homogeneous dynamics, implying the fluctuation theorem symmetry between the forward and backward rate functions. All of these results are subsequently proved for the entropy production in place of the action functional, along with the consequent generalization of the second law of thermodynamics and a derivation of the asymptotic time-averaged entropy production and associated production rate. We conclude with a look beyond periodic driving, considering the case of uniformly continuous and bounded rates.

2 Definitions and Setup

Let $(X_t)_{t \geq 0}$ be a continuous time Markov chain on the probability space (Ω, \mathcal{F}, P) , where $\Omega = D(0, \infty)$ is the space of càdlàg paths (i.e., right-continuous with left limits) over a finite state space, which, without loss of generality, we take to be $\{1, 2, \dots, N\}$, and P is the Markov path measure with initial distribution π . The process generates the right-continuous filtration $\mathcal{F}_t = \bigcap_{s > t} \sigma(X_r, 0 \leq r \leq s)$, representing information known about the process up to and infinitesimally beyond time t , with $\mathcal{F}_t \uparrow \mathcal{F}$, and evolves according to transition rates $k_{ij}(t)$, such that the probability of jumping from state i to j within a time window $[t, t + \tau]$ equals $k_{ij}(t)\tau + o(\tau)$. These rates are assumed to be such that the infinitesimal generator $A(t) = (k_{ij}(t))_{i,j=1}^N$ is continuous $\forall t \geq 0$, T -periodic, and irreducible for some $t^* \in [0, T]$, but whose adjacency graph, representing which states currently communicate, may otherwise change over time and become reducible. We further require that the rates satisfy the so-called dynamic reversibility [27] or ergodic consistency condition [8], whereby $k_{ij}(t) > 0 \iff k_{ji}(t) > 0$, which again ensures that the time-reversal of any realizable trajectory in the forward process is realizable in the backward one. The law of the process, $\mu(i, t) = P(\omega \in \Omega : X_t(\omega) = i)$, satisfies the forward Kolmogorov equation [24]

$$\frac{\partial \mu}{\partial t} = A^*(t)\mu \tag{1}$$

with initial condition $\mu(\cdot, 0) = \pi$, where \cdot^* denotes the adjoint. We take $(\tau_i)_{i \geq 1}$ to be the random jumping times of the process and $\sigma_i = X_{\tau_i}$, with $\sigma_0 = X_0$.

Let us now define the objects we will primarily be concerned with, the entropy production

$$S(t_0, t) = \log \frac{dP_{[t_0, t]}}{dP_{[0, t-t_0]}^B},$$

equal to the logarithmic Radon-Nikodym derivative between the forward and backward path measures (to be discussed shortly), and the action functional

$$W(t_0, t) = S(t_0, t) - \log \frac{\mu(X_{t_0}, t_0)}{\mu(X_t, t)} = \sum_{t_0 < \tau_i < t} \log \frac{k_{\sigma_{i-1}\sigma_i}(\tau_i)}{k_{\sigma_i\sigma_{i-1}}(\tau_i)} \tag{2}$$

representing heat dissipation, where $\log(\mu(X_{t_0}, t_0)/\mu(X_t, t))$ equals the net difference in system entropy between times t_0 and t . ($\mu(X_s, s)$ here should be interpreted as $\mu(i, s)|_{i=X_s(\omega)}$.) Exponential factors representing holding times between jumps have been canceled on the RHS of (2) (see [15, 17]); a full representation of the forward path density can be found in (29) in Sect. 5. For a comprehensive justification of the definitions used above for entropy production and heat dissipation, see [17, 19, 26, 27, 33].

$P_{[t_0, t]}$ above is our Markov path measure restricted to $[t_0, t]$ and $P_{[0, t-t_0]}^B$ is the corresponding backward path measure, defined as follows. Let $P_{[0, t-t_0]}^-$ be a measure obtained from $P_{[t_0, t]}$ by setting its initial distribution $\mu^-(\cdot, 0) = \mu(\cdot, t)$ and its rates $k_{ij}^-(s) = k_{ij}(t - s)$ for $0 \leq s \leq t - t_0$, and let $r(\omega)_s = \lim_{s' \uparrow t-s} \omega_{s'}$ be the path-reversal transformation, with $\omega \in \Omega$ a sample path. We then define $P_{[0, t-t_0]}^B \equiv P_{[0, t-t_0]}^- \circ r$. Note that r is involutive on Ω , preserving the càdlàg property of paths, and that the law of the backward process, $\mu^B(\cdot, s)$, defined for $0 \leq s \leq t - t_0$, is implicitly a function of t due to r . It satisfies the final condition $\mu^B(\cdot, t) = \mu^-(\cdot, 0) = \mu(\cdot, t)$. Note crucially that the backward process is not a Markov process with respect to the filtration $(\mathcal{F}_s)_{0 \leq s \leq t}$. Indeed, $\mu^B(\cdot, 0) = \mu^-(\cdot, t)$, and so the smallest σ -algebra that the event $\{X_0 = i\}$ is measurable with respect to under this process is \mathcal{F}_t . It is, however, Markovian with respect to the backward time filtration $\mathcal{G}_t = \bigcap_{s < t} \sigma(X_r, s \leq r)$.

Reversing both the path and the rates in the backward path measure ensures that the transitions of a reversed path in the P^- -governed process occur under the same local conditions (i.e., instantaneous rates) as those of the original path do in the forward process. Intuitively, if $P_{[0,t]}(\omega)$ is the probability of observing the trajectory ω unfold as time runs forward from 0 to t , then $P_{[0,t]}^B(\omega)$ is the probability of observing that trajectory unfold *in reverse* as time runs *backward* from t to 0, with the filtration \mathcal{G}_t representing the past. From an Archimedian perspective “outside of time” [28], neither direction of time should be preferred a priori. $S(0, t)$ is therefore truly a measure of irreversibility: Thinking of ω as a spacetime curve instead of an oriented trajectory, $S(0, t)(\omega)$ gives the log likelihood of ω being realized in a universe where time runs forward as opposed one in which time runs backward, with the sign of its expectation—dependent on which measure we integrate with respect to—indicating the overall direction of time’s arrow.

It has been argued recently [10], we should note, that the expectation of $S(0, t)$, the mutual entropy $\mathcal{H}(P_{[0,t]}, P_{[0,t]}^B)$, is too sensitive to rare irreversible events to be a useful measure of irreversibility, and that the quantity $A = \frac{1}{2}\mathcal{H}(P_{[0,t]}, \frac{1}{2}(P_{[0,t]} + P_{[0,t]}^B)) + \frac{1}{2}\mathcal{H}(P_{[0,t]}^B, \frac{1}{2}(P_{[0,t]} + P_{[0,t]}^B))$ should be used instead. The “asymmetry” A equals the amount of information gained about the direction of time’s arrow from watching one realization of the process.

The first theorem that we will prove is the existence and differentiability of the free energy function of the action functional $W(0, t)$,

$$c_W(\lambda) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_{\pi,0}(e^{\lambda W(0,t)}), \tag{3}$$

which guarantees a large deviation property for $W(0, t)/t$ with rate function $I_W(z) = \sup_{\lambda \in \mathbb{R}} \{\lambda z - c_W(\lambda)\}$. Here $\mathbb{E}_{\nu,t_0}(\cdot)$ denotes expectation conditioned on the chain having distribution ν at time t_0 . Our second main result is the symmetry relation $c_W(\lambda) = c_W^B(-(1+\lambda))$ between it and the free energy of the action functional under the backward process,

$$c_W^B(\lambda) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_{\mu^{(\cdot,t)},t}^B(e^{\lambda W^B(0,t)}), \tag{4}$$

which reduces to the usual Gallavotti-Cohen symmetry when the generator is a symmetric function of time within each driving period, in which case the backward path measure reduces to the path-reversed measure $P_{[0,t]} \circ r$ considered in [18, 19, 24, 26, 27]. (Note that even in this scenario these authors’ results do not apply, because the dynamics remain time-dependent and non-stationary.) See [32] for a discussion in the context of transient fluctuation theorems.

In the definition (4), the expectation is taken with respect to $P_{[0,t]}^B$, and $W^B(0, t)$ represents the heat dissipated by a process that traces the curve ω backward through time. Since each transition $\sigma_i \rightarrow \sigma_{i-1}$ at time τ_i increases $W^B(0, t)$ by $\log k_{\sigma_i \sigma_{i-1}}(\tau_i)/k_{\sigma_{i-1} \sigma_i}(\tau_i)$, we simply have $W^B(0, t) = -W(0, t)$. Similarly, the entropy production under the backward process $S^B(0, t) = \log dP_{[0,t]}^B/dP_{[0,t]} = -S(0, t)$. The key point is that while W and W^B (resp. S and S^B) are distinct functions, when evaluated under $P_{[0,t]}$ and $P_{[0,t]}^B$, respectively, they represent the *same physical quantity*. For additional discussion of this and the relation between forward and backward measures in general, see Sect. 3.1 of [17].

3 Large Deviations of Heat Dissipation

To prove the existence of the limit (3), we first derive a backward equation for the time-dependent moment generating function $u_\lambda(t_0, t)$, whose elements $u_\lambda(i, t_0, t)$ equal

$\mathbb{E}_{i,t_0}[e^{\lambda W(t_0,t)}]$, where the expectation is conditioned on the process being at state i at time t_0 . For notational economy we define $\Lambda_\lambda(t_0, t) = e^{\lambda W(t_0,t)}$ and $w(i, j, t) = \log(k_{ij}(t)/k_{ji}(t))$. As a general rule we use λ as a subscript in order to emphasize its distinction from time parameters, with the exception of the free energy. Finally, we use the convention that all matrix inequalities, denoted by \preceq , and limits are defined component-wise.

Proposition 3.1 $u_\lambda(t_0, t)$ satisfies the Kolmogorov backward equation

$$\frac{\partial}{\partial t_0} u_\lambda(t_0, t) = -L_\lambda(t_0)u_\lambda(t_0, t) \tag{5}$$

where

$$L_\lambda(t_0)_{i,j} = \begin{cases} k_{ij}(t_0)^{1+\lambda}k_{ji}(t_0)^{-\lambda}, & i \neq j \\ -K_i(t_0), & i = j, \end{cases} \tag{6}$$

and $K_i(t_0) = \sum_{j \neq i}^N k_{ij}(t_0)$ is the escape rate from state i at time t_0 .

Proof Because the jump times of a Markov chain are isolated almost surely and its paths are càdlàg, we interpret the partial derivative in (5) to be the left derivative: $\frac{\partial}{\partial t_0} u_\lambda(t_0, t) = \lim_{h \rightarrow 0^+} \frac{u_\lambda(t_0,t) - u_\lambda(t_0-h,t)}{h}$. Evaluating the second term in the numerator is the crux of the problem. We begin with

$$\begin{aligned} \lim_{h \rightarrow 0^+} u_\lambda(i, t_0 - h, t) &= \lim_{h \rightarrow 0^+} \mathbb{E}_{i,t_0-h}[\Lambda_\lambda(t_0 - h, t)] \\ &= \lim_{h \rightarrow 0^+} \mathbb{E}_{i,t_0-h}[\mathbb{E}_{i,t_0-h}(\Lambda_\lambda(t_0 - h, t_0)\Lambda_\lambda(t_0+, t)|\mathcal{F}_{t_0})] \\ &= \lim_{h \rightarrow 0^+} \mathbb{E}_{i,t_0-h}[\Lambda_\lambda(t_0 - h, t_0)\mathbb{E}_{i,t_0-h}(\Lambda_\lambda(t_0+, t)|\mathcal{F}_{t_0})] \\ &= \lim_{h \rightarrow 0^+} \mathbb{E}_{i,t_0-h}[\Lambda_\lambda(t_0 - h, t_0)\mathbb{E}_{X_{t_0},t_0}(\Lambda_\lambda(t_0+, t))]. \end{aligned}$$

Here the limit t_0+ is taken before the limit $h \downarrow 0$, and in the last line we have used the Markov property inherited by $\Lambda_\lambda(t_0, t)$ from X_t . Since t_0 is not an accumulation point for jumps almost surely and we are operating in the small h limit, at most one jump can occur between $t_0 - h$ and t_0 . The last line therefore becomes

$$\begin{aligned} &\lim_{h \rightarrow 0^+} \mathbb{E}_{i,t_0-h} [e^{\lambda w(i, X_{t_0}, t_0)} \mathbb{E}_{X_{t_0},t_0} (\Lambda_\lambda(t_0+, t))] \\ &= \lim_{h \rightarrow 0^+} \sum_{j=1}^N \mathbb{E}_{i,t_0-h} [e^{\lambda w(i, X_{t_0}, t_0)} \mathbb{E}_{X_{t_0},t_0} (\Lambda_\lambda(t_0+, t)) 1_{X_{t_0}=j}] \\ &= \lim_{h \rightarrow 0^+} \sum_{j=1}^N e^{\lambda w(i, j, t_0)} \mathbb{E}_{i,t_0-h} [\mathbb{E}_{j,t_0} (\Lambda_\lambda(t_0+, t)) 1_{X_{t_0}=j}], \end{aligned}$$

with 1_A denoting the indicator of the event A . Regardless of whether a jump occurs at t_0 (that is, $X_{t_0} = j \neq i$), no jumps may occur during a sufficiently small period after t_0 . The inner expectation therefore equals $u_\lambda(j, t_0, t)$ and our limit is

$$\lim_{h \rightarrow 0^+} \sum_{j=1}^N e^{\lambda w(i, j, t_0)} u_\lambda(j, t_0, t) \mathbb{E}_{i,t_0-h} [1_{X_{t_0}=j}]$$

$$= \lim_{h \rightarrow 0^+} \sum_{j=1}^N e^{\lambda w(i,j,t_0)} u_\lambda(j, t_0, t) p(i, j, t_0 - h, t_0), \tag{7}$$

where $p(i, j, t_0 - h, t)$ is the transition probability of being at j at time t_0 given having been at i at $t_0 - h$. For $i \neq j$ this equals $k_{ij}(t_0 - h)h + o(h)$ and, for $i = j$, it equals $1 - K_i(t_0 - h)h + o(h)$, with K_i defined in the statement of the Proposition. We therefore have, by continuity of the k_{ij} ,

$$\begin{aligned} \lim_{h \rightarrow 0^+} u_\lambda(i, t_0 - h, t) &= \lim_{h \rightarrow 0^+} \sum_{j \neq i}^N (k_{ij}(t_0)h + o(h)) e^{\lambda w(i,j,t_0)} u_\lambda(j, t_0, t) \\ &\quad + (1 - K_i(t_0)h + o(h)) u_\lambda(i, t_0, t) \end{aligned}$$

and the component-wise derivative becomes

$$\begin{aligned} \frac{\partial}{\partial t_0} u_\lambda(i, t_0, t) &= \lim_{h \rightarrow 0^+} \frac{1}{h} [u_\lambda(i, t_0, t) - u_\lambda(i, t_0 - h, t)] \\ &= - \left[\sum_{j \neq i}^N k_{ij}(t_0)^{1+\lambda} k_{ji}(t_0)^{-\lambda} u_\lambda(j, t_0, t) - K_i(t_0) u_\lambda(i, t_0, t) \right]. \end{aligned}$$

Collecting components into a vector equation, we obtain (5). □

Before stating the first main theorem, we introduce a concept from quantum field theory that will be useful. For a family of operators $\{\mathcal{O}(s)\}_{s \in \mathbb{R}}$, define the time-ordering operator \mathcal{T} [36] by $\mathcal{T} \left[\prod_{j=0}^n \mathcal{O}(s_j) \right] = \mathcal{O}(s_n) \cdots \mathcal{O}(s_1) \mathcal{O}(s_0)$, where $s_0 < \cdots < s_n$ is any finite sequence in \mathbb{R} . Now, given $t_0 < t_1$, we define the time-ordered exponential

$$\begin{aligned} \mathcal{T} \left(\exp \int_{t_0}^{t_1} \mathcal{O}(s) ds \right) &= \lim_{n \rightarrow \infty} \exp \left(\frac{t_1 - t_0}{n} \mathcal{O}(s_{n-1}^n) \right) \cdots \exp \left(\frac{t_1 - t_0}{n} \mathcal{O}(s_1^n) \right) \\ &\quad \times \exp \left(\frac{t_1 - t_0}{n} \mathcal{O}(s_0^n) \right) \end{aligned}$$

when the limit exists, where $\{s_j^n = t_0 + j(t_1 - t_0)/n\}_{j=0}^n$ is a sequence of uniform partitions of $[t_0, t_1]$.

Recall that a matrix is called *quasipositive* if its entries are nonnegative, strictly on the diagonal. Our results rely on the fact that the time-ordered exponential operator, like the usual exponential operator, transforms quasipositivity into strict positivity. The proof requires some clever bounding arguments.

Lemma 3.2 *Let $0 \leq t_0 < t_1$ and $L : [0, \infty) \rightarrow M_{N \times N}(\mathbb{R})$ be a continuous function whose values are quasipositive $\forall t \geq 0$ and irreducible for some $t^* \in (t_0, t_1)$. Then the matrix $\mathcal{T}(\exp \int_{t_0}^{t_1} L(t) dt)$ has strictly positive entries.*

Proof Without loss of generality, we take $t_0 = 0$, the length of the interval being the only relevant factor. By uniform continuity of $L(t)$ on $[0, t_1]$, choose $\sup_{t \in [0, t_1], 1 \leq i \leq N} L(t)_{ii}^- < M < \infty$, where the superscript denotes the negative part of the number, so that $E(t) =$

$L(t) + MI$ has strictly positive diagonal entries and nonnegative off-diagonal ones, $\forall t \in [0, t_1]$. Employing the partitions $\{s_j^n\}$ defined above, we have

$$\begin{aligned} & \exp\left(\frac{t_1}{n}L(s_{n-1}^n)\right) \cdots \exp\left(\frac{t_1}{n}L(s_0^n)\right) \\ &= \exp\left(\frac{t_1}{n}L(s_{n-1}^n) + \frac{Mt_1}{n}I - \frac{Mt_1}{n}I\right) \cdots \exp\left(\frac{t_1}{n}L(s_0^n) + \frac{Mt_1}{n}I - \frac{Mt_1}{n}I\right) \\ &= e^{-Mt_1} \exp\left(\frac{t_1}{n}E(s_{n-1}^n)\right) \cdots \exp\left(\frac{t_1}{n}E(s_0^n)\right) \\ &\geq e^{-Mt_1} \left(I + \frac{t_1}{n}E(s_{n-1}^n)\right) \cdots \left(I + \frac{t_1}{n}E(s_0^n)\right). \end{aligned} \tag{8}$$

By irreducibility of $L(t^*)$, let $0 < \alpha < \min_{1 \leq i, j \leq N} (E(t^*)^N)_{i, j}$, and define E_α to be the $N \times N$ matrix whose entries are all α . By continuity of $E(t)$, choose δ small enough so that $\forall t^* < u_1, \dots, u_N < t^* + \delta, E(u_1) \cdots E(u_N) \geq E_\alpha$. Finally, let $k = \lceil nt^*/t_1 \rceil$ and $\ell = \lfloor n(t^* + \delta)/t_1 \rfloor$ denote the index of the first partition point after t^* and the last one before $t^* + \delta$, respectively, so that $\ell - k \geq n\delta/t_1 - 2$. We can then rewrite (8) as

$$\begin{aligned} & e^{-Ms_1} \left(I + \frac{t_1}{n}E(s_{n-1}^n)\right) \cdots \left(I + \frac{t_1}{n}E(s_{\ell+1}^n)\right) \left(I + \frac{t_1}{n}E(s_\ell^n)\right) \cdots \left(I + \frac{t_1}{n}E(s_k^n)\right) \\ & \quad \times \left(I + \frac{t_1}{n}E(s_{k-1}^n)\right) \cdots \left(I + \frac{t_1}{n}E(s_0^n)\right) \\ & \geq e^{-Mt_1} \left(I + \frac{t_1}{n}E(s_\ell^n)\right) \cdots \left(I + \frac{t_1}{n}E(s_k^n)\right) \geq e^{-Mt_1} \binom{\ell - k}{N} \left(\frac{t_1}{n}\right)^N E_\alpha \\ & \geq e^{-Mt_1} \frac{(n\delta/t_1 - 2) \cdots (n\delta/t_1 - (N + 1))}{(n/t_1)^N} \frac{1}{N!} E_\alpha \xrightarrow{n \rightarrow \infty} e^{-Mt_1} \frac{\delta^N}{N!} E_\alpha. \end{aligned}$$

Therefore, choosing $0 < \beta < e^{-Mt_1} \delta^N \alpha / N!$, we have that the entries of $\mathcal{T}(\exp \int_0^{t_1} L(t) dt)$ are bounded below by β . □

Note that by the irreducibility and ergodic consistency conditions on $A(t)$, $L_\lambda(t)$ is irreducible at t^* , and, hence, the lemma applies to it. We recall that our assumptions on the generator $A(t)$ are sufficiently weak that the rates $k_{ij}(t)$ may drop down to zero, thereby changing the structure of its adjacency graph and possibly rendering it reducible. Irreducibility of $A(t)$ over an arbitrarily small interval is sufficient, nevertheless, for all states of the chain to communicate after t^* .

Theorem 3.3 *Given the periodicity, continuity, irreducibility and ergodic consistency conditions on the generator $A(t)$, the free energy $c_W(\lambda)$ exists, is continuously differentiable $\forall \lambda \in \mathbb{R}$, and is independent of π .*

Proof Rewriting the free energy as $\lim_{t \rightarrow \infty} \frac{1}{t} \log \langle \pi, u_\lambda(0, t) \rangle$, where $\langle \cdot \rangle$ denotes the inner-product on \mathbb{R}^N , our strategy is to represent $u_\lambda(0, t)$ in terms of the Floquet fundamental solution to (5), whose associated flow operator we will show has strictly positive entries and

whose periodic component can be bounded so as not to affect the asymptotics. The Perron-Frobenius theorem will then allow us to identify the free energy as the principal Floquet eigenvalue of the fundamental solution, which depends smoothly on λ .

We begin by noticing that in our backward equation, $u_\lambda(t_0, t)$ evolves according to the matrix $-L_\lambda(t_0)$, whose off-diagonal terms are non-positive and whose corresponding flow will therefore not have exclusively positive entries. Instead of integrating forward, we therefore elect to change variables and integrate *back* to $u_\lambda(0, t)$ from the known value $u_\lambda(t, t) \equiv (1 \ 1 \cdots 1)^*$. To wit, let $\tau = t - t_0$, and define $\tilde{u}_\lambda(\tau, t) = u_\lambda(t - \tau, t) = u_\lambda(t_0, t)$ and similarly $\tilde{L}_\lambda(\tau) = L_\lambda(t_0)$. Then

$$\frac{\partial}{\partial \tau} \tilde{u}_\lambda(\tau, t) = -\frac{\partial}{\partial t_0} u_\lambda(t_0, t) = L_\lambda(t_0)u_\lambda(t_0, t) = \tilde{L}_\lambda(\tau)\tilde{u}_\lambda(\tau, t). \tag{9}$$

Recall that the fundamental solution to (9) with initial condition $\tilde{u}_\lambda(0, t) = (1 \ 1 \cdots 1)^*$ is a square matrix whose N columns are N linearly independent solutions to the ODE. The Floquet theorem [41] guarantees that by periodicity of $\tilde{L}_\lambda(\tau)$, this fundamental solution has the form $\Phi_\lambda(\tau, t) = e^{\tau H_\lambda(t)} P_\lambda(\tau, t)$, where P_λ is T -periodic and continuous in τ and $H_\lambda(t)$ is complex. We choose in particular the principle fundamental solution, which satisfies $\Phi_\lambda(0, t) = I$, so that $\tilde{u}_\lambda(\tau, t) = \Phi_\lambda(\tau, t)\tilde{u}_\lambda(0, t)$. This theorem is derived by first noting that by periodicity of \tilde{L}_λ , if $\Phi_\lambda(\tau, t)$ is a fundamental solution of (9), then so is $\Phi_\lambda(\tau + T, t)$. This implies that the two matrices are linearly dependent, and hence there exists a nonsingular matrix $C(t)$ such that $\Phi_\lambda(\tau + T, t) = C(t)\Phi_\lambda(\tau, t)$. $H_\lambda(t)$ is defined as $\frac{1}{T} \log C(t)$, since nonsingular matrices always possess a logarithm. We may therefore identify $e^{T H_\lambda(t)}$ as the flow operator that evolves solutions T units of time into the future, which is *independent* of the current time τ . From (9), it can then be represented as

$$e^{T H_\lambda(t)} = T \left(\exp \int_\tau^{\tau+T} \tilde{L}_\lambda(s) ds \right), \quad \forall 0 \leq \tau \leq t - T, \tag{10}$$

which, by the lemma, has strictly positive entries. A second dividend of this representation is that because the RHS does not depend on t , neither does H_λ .

By the Perron-Frobenius theorem, $e^{T H_\lambda}$ possesses a nondegenerate positive, maximum eigenvalue $e^{T \vartheta(\lambda)}$ and a positive eigenvector v_λ spanning its corresponding one-dimensional eigenspace. This makes $\vartheta(\lambda)$ the nondegenerate principal Floquet eigenvalue of the fundamental solution Φ_λ , which, as a simple root of the characteristic equation of H_λ , is continuously differentiable with respect to λ by the implicit function theorem [23]. Evaluating the free energy function,

$$\begin{aligned} c_W(\lambda) &= \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_{\pi, 0} (e^{\lambda W(0, t)}) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \langle \pi, u_\lambda(0, t) \rangle \\ &= \lim_{t \rightarrow \infty} \frac{1}{t} \log \langle \pi, \tilde{u}_\lambda(t, t) \rangle = \lim_{t \rightarrow \infty} \frac{1}{t} \log \langle \pi, e^{t H_\lambda} P_\lambda(t, t) (1 \ 1 \cdots 1)^* \rangle \\ &= \lim_{t \rightarrow \infty} \frac{1}{t} \log \left(\langle \pi, v_\lambda \rangle e^{t \vartheta(\lambda)} \langle v_\lambda, P_\lambda(t, t) (1 \ 1 \cdots 1)^* \rangle \right) \\ &= \vartheta(\lambda) + \lim_{t \rightarrow \infty} \frac{1}{t} \log \langle v_\lambda, P_\lambda(t, t) (1 \ 1 \cdots 1)^* \rangle \end{aligned} \tag{11}$$

where we have used the uniqueness of the maximal eigenvalue and then the positivity of v_λ to eliminate the dependence on π .

Our final task is to show the remaining limit vanishes by bounding the quadratic form $Q(t) = \langle v_\lambda, P_\lambda(t, t)(1 \ 1 \cdots 1)^* \rangle$ between positive numbers, uniformly in t . To achieve this, we again exploit the time-ordered exponential representation of the flow. Taking $t = mT + r$, where $0 \leq r < T$, by the factoring property of such operators [36]

$$\begin{aligned} e^{mTH_\lambda} e^{rH_\lambda} P_\lambda(t, t) &= \mathcal{T} \left(\exp \int_0^t \tilde{L}_\lambda(s) ds \right) = \mathcal{T} \left(\exp \int_r^t \tilde{L}_\lambda(s) ds \right) \mathcal{T} \left(\exp \int_0^r \tilde{L}_\lambda(s) ds \right) \\ &= e^{mTH_\lambda} \mathcal{T} \left(\exp \int_0^r \tilde{L}_\lambda(s) ds \right). \end{aligned}$$

Multiplying by $e^{-rH_\lambda} e^{-mTH_\lambda}$ on the left, we have

$$P_\lambda(t, t) = e^{-rH_\lambda} \mathcal{T} \left(\exp \int_0^r \tilde{L}_\lambda(s) ds \right). \tag{12}$$

The presence of r but not t on the RHS reveals that $\tilde{P}_\lambda(t) \equiv P_\lambda(t, t)$ is T -periodic, and the form of (12) ensures that it is uniformly continuous on its periodic domain. We may conclude from this that the matrix norm of $P_\lambda(t, t)$ is bounded in t , and so the quadratic form is bounded above.

Because we may have $r < t^*$, the time at which the generator is irreducible, the lemma does not guarantee for us that $\mathcal{T}(\exp \int_0^r \tilde{L}_\lambda(s) ds)$ has strictly positive entries. However, it is straightforward to see by adding and subtracting $M > \sup_{s \in [0, r], 1 \leq i \leq N} \tilde{L}_\lambda(s)_{ii}$ times the identity from $\tilde{L}_\lambda(s)$ that this matrix exponential does have nonnegative entries, strictly on the diagonal. From this we may conclude that

$$w_\lambda \equiv \mathcal{T} \left(\exp \int_0^r \tilde{L}_\lambda(s) ds \right) (1 \ 1 \cdots 1)^* \geq 0$$

with $w_\lambda \neq 0$. Taking $R_{v_\lambda^\perp}$ to be the projection operator onto the orthogonal subspace of v_λ , by (12),

$$\begin{aligned} Q(t) &= \langle v_\lambda, P_\lambda(t, t)(1 \ 1 \cdots 1)^* \rangle \\ &= \langle v_\lambda, e^{-rH_\lambda} \mathcal{T} \left(\exp \int_0^r \tilde{L}_\lambda(s) ds \right) (1 \ 1 \cdots 1)^* \rangle \\ &= \langle v_\lambda, e^{-rH_\lambda} w_\lambda \rangle = \langle v_\lambda, e^{-rH_\lambda} (\langle v_\lambda, w_\lambda \rangle v_\lambda + R_{v_\lambda^\perp} w_\lambda) \rangle = e^{-r\theta(\lambda)} \langle v_\lambda, w_\lambda \rangle^2 > 0. \end{aligned}$$

This is our uniform lower bound. □

Remark 3.4 One novelty of the proof of Theorem 3.3 vis-à-vis the proofs of large deviation principles by Lebowitz and Spohn [24], Jiang et al. [18] and others is that it does not rely on the existence of an ergodic distribution for the dynamics. Indeed, no stationary distribution exists. By (1), $\mu(\cdot, t)$ evolves according to the periodic matrix $A^*(t)$, and so by the Floquet theorem can be represented as $\mu(\cdot, t) = e^{tB} P(t) \pi$, where $P(t)$ is T -periodic. Because $A^*(t)$ is quasipositive always and irreducible at t^* , by Lemma 3.2 and an argument analogous to that in the proof of Theorem 3.3, e^{tB} has strictly positive entries and thus a Perron-Frobenius eigenvalue. This eigenvalue must be 1, otherwise probability conservation would

be violated. Taking v to be the corresponding positive eigenvector and R_{v^\perp} the projection onto the subspace of \mathbb{R}^N orthogonal to it,

$$\mu(\cdot, t) = e^{tB} P(t)\pi = \langle P(t)\pi, v \rangle v + e^{tB} R_{v^\perp} P(t)\pi \sim \langle P(t)\pi, v \rangle v. \tag{13}$$

We see that the asymptotic limit of *any* initial distribution π is T -periodic through $P(t)$, and therefore that no stationary distribution exists. As proved in Theorem 5.2, this asymptotic limit takes the place of the usual stationary distribution in the expression for the instantaneous entropy production rate.

Corollary 3.5 *The time-averaged action functional $W(0, t)/t$ satisfies a large deviation principle with continuous and strictly convex good rate function*

$$I_W(z) = \sup_{\lambda \in \mathbb{R}} \{\lambda z - c_W(\lambda)\}. \tag{14}$$

Proof Existence and differentiability of $c_W(\lambda)$ everywhere imply, by the Gärtner-Ellis theorem [5], that $W(0, t)/t$ satisfies a large deviation principle with rate function $I_W(z) = \sup_{\lambda \in \mathbb{R}} \{\lambda z - c_W(\lambda)\}$. Strict convexity of $I_W(z)$ can be deduced by contradiction. If the convex envelope of $I_W(z)$ (the supremum of all convex functions minorizing it) contained an interval of strict linearity with slope λ_0 , then the free energy’s derivative would jump at λ_0 by Legendre duality, violating its continuity (see [38] for an intuitive discussion).

Continuity requires a more detailed argument. Differentiability of $c_W(\lambda)$ implies that (14) can be computed from calculus: $I_W(z) = \lambda_z z - c_W(\lambda_z)$, where λ_z is defined implicitly through $c'_W(\lambda_z) = z$. As the free energy may not be strictly convex, however, λ_z is not necessarily unique. The set $\{\lambda_z \in \mathbb{R} : c'_W(\lambda_z) = z\}$ is in fact a closed interval $[\lambda_z^{\min}, \lambda_z^{\max}]$, which reduces to a single point only for z corresponding to non-linear portions of $c_W(\lambda)$. The RHS of (14) achieves its maximum at every λ_z in this interval, including, in particular, at the endpoints.

Let us now consider a sequence z_n converging to z . We then have $\liminf_{n \rightarrow \infty} \lambda_{z_n} \geq \lambda_z^{\min}$ and $\limsup_{n \rightarrow \infty} \lambda_{z_n} \leq \lambda_z^{\max}$, which, by continuity of $c_W(\lambda)$, implies that

$$\begin{aligned} \limsup_{n \rightarrow \infty} I_W(z_n) &= \limsup_{n \rightarrow \infty} \lambda_{z_n} z_n - c_W(\lambda_{z_n}) \leq I_W(z) \\ &\leq \liminf_{n \rightarrow \infty} \lambda_{z_n} z_n - c_W(\lambda_{z_n}) = \liminf_{n \rightarrow \infty} I_W(z_n), \end{aligned}$$

or $\lim_{n \rightarrow \infty} I_W(z_n) = I_W(z)$.

Finally, recall that $I_W(z)$ is called a *good rate function* if all sets of the form $\{z \in \mathbb{R} : I_W(z) \leq a\}$ are compact. This is implied directly by convexity and continuity. \square

Remark 3.6 Differentiability of the free energy is crucial in the proof of the corollary. Without it, $I_W(z)$ may be non-convex, with the Legendre-Fenchel transform of $c_W(\lambda)$ yielding only its convex envelope. In such situations there exists no general method for accessing the rate function. This breakdown in Legendre duality has been proved [7] to be a necessary and sufficient condition for the nonequivalence of the canonical and microcanonical ensembles in statistical mechanics, whereby mean energies equal to slopes not present in the convex free energy profile (due to a nonanalyticity) cannot be realized in the canonical ensemble. Even given continuous differentiability of the free energy, however, we cannot improve upon the continuity result for the rate function to conclude differentiability. The reason is that the free energy may contain linear parts, each of which creates a nonanalyticity in the

rate function. This is why we speak of the two as being related by the Legendre-Fenchel transform instead of the better known Legendre transform, which is defined only between differentiable functions.

We now prove our primary result for the action functional.

Theorem 3.7 *The free energy $c_W^B(\lambda)$ exists, is continuously differentiable $\forall \lambda \in \mathbb{R}$ and satisfies, together with $c_W(\lambda)$, the symmetry relation*

$$c_W(\lambda) = c_W^B(-(1 + \lambda)), \tag{15}$$

implying the fluctuation theorem

$$I_W(z) - I_W^B(-z) = -z, \tag{16}$$

where I_W^B is the continuous and strictly convex good rate function of the action functional under the backward process.

Proof Obtaining an explicit representation of the free energy $c_W^B(\lambda)$ is simply a matter of bootstrapping from the work we did in Proposition 3.1 and Theorem 3.3. It is implied by the discussion in Sect. 2 that if we reverse time and the orientation of the driving rates $k_{ij}(t)$, then the forward process becomes the backward process up to boundary conditions. In particular, the backward action functional $W^B(t_0, t)(\omega)$ equals the heat accumulated by traversing the path ω backward from time $t - t_0$ to time 0. This immediately implies

$$\frac{\partial}{\partial t_0} u_\lambda^B(t_0, t) = -L_\lambda(t - t_0) u_\lambda^B(t_0, t)$$

for the moment generating function $u_\lambda^B(t_0, t)$, with components $u_\lambda^B(i, t_0, t) = \mathbb{E}_{i, t_0}^B [e^{\lambda W^B(t_0, t)}]$. Intuitively, increasing t_0 still decreases the moment generating function, only in the backward process the increments that are lost correspond to jumps at $t - t_0$. Were we to recapitulate the proof of Proposition 3.1, the filtration \mathcal{F}_t used in the conditional expectations would be replaced by \mathcal{G}_t , representing equivalently the future of the forward process and the past of the backward one. Making the change of variables $\tau = t - t_0$, with $\tilde{u}_\lambda^B(\tau, t) = u_\lambda^B(t_0, t)$, we obtain

$$\frac{\partial}{\partial \tau} \tilde{u}_\lambda^B(\tau, t) = \mathbf{L}_\lambda(\tau) \tilde{u}_\lambda^B(\tau, t). \tag{17}$$

The solution to this equation can be represented as $\tilde{u}_\lambda^B(\tau, t) = e^{\tau H_\lambda^B} P_\lambda^B(\tau, t) (1 \ 1 \ \dots \ 1)^*$ for some T -periodic P_λ^B and real H_λ^B , whose matrix exponential has strictly positive entries. Writing the free energy as $c_W^B(\lambda) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \langle \mu(\cdot, t), \tilde{u}_\lambda^B(t, t) \rangle$, by a chain of equalities similar to (11) and the ensuing argument, $c_W^B(\lambda)$ equals the continuously differentiable Perron-Frobenius eigenvalue of H_λ^B . The desired properties for its rate function $I_W^B(z)$ then follow from the arguments in Corollary 3.5.

We now use the symmetry relation $L_{-(1+\lambda)}^* = L_\lambda$ (and $\tilde{L}_{-(1+\lambda)}^* = \tilde{L}_\lambda$) observed from (6) to prove that

$$(H_{-(1+\lambda)}^B)^* = H_\lambda. \tag{18}$$

Indeed, as $e^{TH_\lambda^B}$ is a flow operator, mapping solutions of (17) T units of time forward (in the direction of increasing τ), irrespective of the current time, it has the representation

$$e^{TH_\lambda^B} = \mathcal{T} \left(\exp \int_\tau^{\tau+T} L_\lambda(s) ds \right), \quad \forall 0 \leq \tau \leq t - T.$$

Fixing one such τ and letting $\{s_j^n\}_{0 \leq j \leq n}$ be the sequence of uniform partitions of $[\tau, \tau + T]$ with mesh size T/n , we then have

$$\begin{aligned} e^{T(H_{-(1+\lambda)}^B)^*} &= \left(e^{TH_{-(1+\lambda)}^B} \right)^* \\ &= \lim_{n \rightarrow \infty} \left\{ \exp \left(\frac{T}{n} L_{-(1+\lambda)}(s_{n-1}^n) \right) \cdots \exp \left(\frac{T}{n} L_{-(1+\lambda)}(s_1^n) \right) \exp \left(\frac{T}{n} L_{-(1+\lambda)}(s_0^n) \right) \right\}^* \\ &= \lim_{n \rightarrow \infty} \exp \left(\frac{T}{n} L_\lambda(s_0^n) \right) \exp \left(\frac{T}{n} L_\lambda(s_1^n) \right) \cdots \exp \left(\frac{T}{n} L_\lambda(s_{n-1}^n) \right) \\ &= \lim_{n \rightarrow \infty} \exp \left(\frac{T}{n} \tilde{L}_\lambda(t - s_0^n) \right) \exp \left(\frac{T}{n} \tilde{L}_\lambda(t - s_1^n) \right) \cdots \exp \left(\frac{T}{n} \tilde{L}_\lambda(t - s_{n-1}^n) \right) \\ &= \mathcal{T} \left(\exp \int_{t-(\tau+T)}^{t-\tau} \tilde{L}_\lambda(s) ds \right) = e^{TH_\lambda}, \end{aligned}$$

where the final equality is justified by the fact that $(t - \tau) - [t - (\tau + T)] = T$. We were also able to interchange the limit and adjoint operators because a sequence of matrices A_n converges component-wise to A iff A_n^* converges to A^* . By the uniqueness of matrix logarithms, (18) holds.

To prove (15), let $H_\lambda v_\lambda = c_W(\lambda) v_\lambda$, where $c_W(\lambda)$ is the maximal eigenvalue of H_λ and v_λ is its corresponding positive eigenvector. By similarity of matrices to their adjoints, $c_W(\lambda)$ is also the maximal eigenvalue of H_λ^* and, hence, of $H_{-(1+\lambda)}^B$, via (18). But we have already shown $c_W^B(-(1 + \lambda))$ to be maximal for $H_{-(1+\lambda)}^B$, and so by uniqueness of the Perron Frobenius eigenvalue, $c_W(\lambda) = c_W^B(-(1 + \lambda))$.

The fluctuation theorem (16) now follows from the Gärtner-Ellis theorem:

$$\begin{aligned} I_W(z) &= \sup_{\lambda \in \mathbb{R}} \{ \lambda z - c_W(\lambda) \} = \sup_{\lambda \in \mathbb{R}} \{ \lambda z - c_W^B(-(1 + \lambda)) \} \\ &= \sup_{\lambda \in \mathbb{R}} \{ -(1 + \lambda)z - c_W^B(\lambda) \} = I_W^B(-z) - z. \end{aligned} \tag{19}$$

□

Remark 3.8 The careful reader will notice that the λ dependence in the relation (15) that we obtain matches that in the free energy symmetries of [18] and [16], but not [24]. The reason is one of convention. The scaled cumulant generating function $e(\lambda) = \lim_{t \rightarrow \infty} -\frac{1}{t} \log \langle e^{-\lambda W(t)} \rangle$ in [24] satisfying the symmetry $e(\lambda) = e(1 - \lambda)$ is the *canonical* free energy function [38], familiar from statistical mechanics when t is the particle number, $W(t)$ the Hamiltonian of a configuration and λ the inverse temperature, but distinct from our free energy function (3) by the minus signs. It nevertheless gives rise to the same symmetry for the rate function as ours (in the special case of homogeneous dynamics) via the Legendre-like identity $I_W(z) = \sup_{\lambda \in \mathbb{R}} \{ e(\lambda) - \lambda z \}$ [24]. For an in depth discussion of the relationship between the physics and mathematical notions of entropy and free energy, see [38].

Remark 3.9 When the generator $A(t)$ is symmetric within each driving period, $I_W^B = I_W$ and our fluctuation theorem (16) reduces to the usual $I_W(z) - I_W(-z) = -z$, which says that the odd part of I_W is linear with slope $-1/2$. When time-inhomogeneities are involved, however, the fluctuation theorem no longer represents an internal symmetry of a single rate function. If we interpret it as a formula for computing the large deviations of heat dissipation in the forward process in terms of those in the backward process, then the invariance of (16) with respect to the joint transformation $I_W^B \leftrightarrow I_W, z \leftrightarrow -z$, which amounts to reversing the roles of the forward and backward process (recall that $W^B(0, t) = -W(0, t)$), indicates an invariance in the large deviations of heat dissipation with respect to time reversal.

4 Large Deviations of Entropy Production

We now prove results analogous to the ones above, but with the action functional $W(0, t)$ replaced by the entropy production $S(0, t)$ of the stochastic particle, whose free energies under the forward and backward processes are

$$c_S(\lambda) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_{\pi,0}(e^{\lambda S(0,t)}) \quad \text{and} \quad c_S^B(\lambda) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_{\mu(\cdot,t)}^B(e^{\lambda S^B(0,t)}). \tag{20}$$

Because the proofs are very similar, we cover only the modifications that must be made.

Theorem 4.1 *Given the periodicity, continuity, irreducibility and ergodic consistency conditions on $A(t)$, the free energy $c_S(\lambda)$ exists, is continuously differentiable $\forall \lambda \in \mathbb{R}$, and is independent of π . The time-averaged entropy production $S(0, t)/t$ therefore satisfies a large deviation principle with continuous and strictly convex good rate function $I_S(z) = \sup_{\lambda \in \mathbb{R}} \{\lambda z - c_S(\lambda)\}$.*

Proof Let $v_\lambda(t_0, t)$ denote the moment generating function of $S(t_0, t)$, with elements $v_\lambda(i, t_0, t) = \mathbb{E}_{i,t_0}[e^{\lambda S(t_0,t)}]$. The derivation of the backward equation we prove for v_λ begins identically to that in Proposition 3.1, except that we replace $w(i, j, t)$ with

$$s(i, j, t_0, t) = \log \frac{k_{ij}(t_0)}{k_{ji}(t_0)} + \log \frac{\mu(i, t_0)}{\mu(j, t)}.$$

These new increments incorporate the boundary terms so that the entropy production becomes an additive process like the action functional, with the representation

$$S(t_0, t) = \lim_{\|\mathcal{P}\| \rightarrow 0} \sum_{i=0}^n s(X_{t_i}, X_{t_{i+1}}, t_i, t_{i+1}),$$

where $\mathcal{P} = \{t_i\}_{i=0}^n$ is a partition of $[t_0, t]$. Starting from the expression (7), we have

$$\begin{aligned} \lim_{h \rightarrow 0+} v_\lambda(i, t_0 - h, t) &= \lim_{h \rightarrow 0+} \sum_{j=1}^N e^{\lambda s(i,j,t_0-h,t_0)} v_\lambda(j, t_0, t) p(i, j, t_0 - h, t_0) \\ &= \lim_{h \rightarrow 0+} \sum_{j \neq i}^N \left(\frac{k_{ij}(t_0 - h)\mu(i, t_0 - h)}{k_{ji}(t_0 - h)\mu(j, t_0)} \right)^\lambda (k_{ij}(t_0)h + o(h)) v_\lambda(j, t_0, t) \end{aligned}$$

$$+ \left(\frac{\mu(i, t_0 - h)}{\mu(i, t_0)} \right)^\lambda (1 - K_i(t_0)h + o(h)) v_\lambda(i, t_0, t).$$

By the forward evolution equation (1) for μ ,

$$\begin{aligned} \left(\frac{\mu(i, t_0 - h)}{\mu(i, t_0)} \right)^\lambda &= 1 - h\lambda \frac{\partial_t \mu(i, t_0)}{\mu(i, t_0)} + o(h) \\ &= 1 - h\lambda \sum_{j \neq i}^N k_{ji}(t_0) \frac{\mu(j, t_0)}{\mu(i, t_0)} + h\lambda K_i(t_0) + o(h), \end{aligned}$$

which we insert into the previous expression to obtain

$$\begin{aligned} \lim_{h \rightarrow 0^+} v_\lambda(i, t_0 - h, t) &= \lim_{h \rightarrow 0^+} \sum_{j \neq i}^N \left(\frac{k_{ij}(t_0 - h)\mu(i, t_0 - h)}{k_{ji}(t_0 - h)\mu(j, t_0)} \right)^\lambda (k_{ij}(t_0)h + o(h)) v_\lambda(j, t_0, t) \\ &\quad + (1 - K_i(t_0)h + o(h)) v_\lambda(i, t_0, t) \\ &\quad - v_\lambda(i, t_0, t) h\lambda \sum_{j \neq i}^N k_{ji}(t_0) \frac{\mu(j, t_0)}{\mu(i, t_0)} + h\lambda K_i(t_0) v_\lambda(i, t_0, t). \end{aligned}$$

Our component-wise derivative for $v_\lambda(i, t_0, t)$ then becomes

$$\begin{aligned} \frac{\partial}{\partial t_0} v_\lambda(i, t_0, t) &= \lim_{h \rightarrow 0^+} \frac{1}{h} [v_\lambda(i, t_0, t) - v_\lambda(i, t_0 - h, t)] \\ &= - \left[\sum_{j \neq i}^N k_{ij}(t_0)^{1+\lambda} k_{ji}(t_0)^{-\lambda} \left(\frac{\mu(i, t_0)}{\mu(j, t_0)} \right)^\lambda v_\lambda(j, t_0, t) \right. \\ &\quad \left. + \left(\lambda \sum_{j \neq i}^N k_{ji}(t_0) \frac{\mu(j, t_0)}{\mu(i, t_0)} + (1 - \lambda) K_i(t_0) \right) v_\lambda(i, t_0, t) \right], \end{aligned} \tag{21}$$

which yields the linear backward equation $\frac{\partial}{\partial t_0} v_\lambda(t_0, t) = -M_\lambda(t_0) v_\lambda(t_0, t)$, where

$$M_\lambda(t_0)_{ij} = \begin{cases} k_{ij}(t_0)^{1+\lambda} k_{ji}(t_0)^{-\lambda} \left(\frac{\mu(i, t_0)}{\mu(j, t_0)} \right)^\lambda, & i \neq j \\ \lambda \sum_{\ell \neq i}^N k_{\ell i}(t_0) \frac{\mu(\ell, t_0)}{\mu(i, t_0)} + (1 - \lambda) K_i(t_0), & i = j. \end{cases} \tag{22}$$

The existence of the off-diagonal terms of $M_\lambda(t_0)$ for $t_0 < t^*$ appears problematic here, because we are dividing by $\mu(j, t_0)$, which may be zero. However, in order for this to be a problem, we must have $k_{ij}(t_0)$ and $\mu(i, t_0)$ both positive. By continuity of these quantities, they must have been positive on $[t_0 - \delta, t_0]$ for some small δ , ensuring that $\mu(j, t_0) > 0$. The same argument justifies the finiteness of the diagonal terms.

The rest of the proof is identical to that of Theorem 3.3 and Corollary 3.5, where we make the change of variables $\tau = t - t_0$ and use the Floquet theorem to factor the new solution $\tilde{v}_\lambda(\tau, t)$ into a periodic matrix and an exponential flow matrix with strictly positive entries. The free energy $c_S(\lambda)$ is then the principal Floquet eigenvalue of the latter, which

by uniqueness is a continuously differentiable function of λ , implying a large deviation principle with continuous and strictly convex good rate function equal to the Legendre-Fenchel transform of $c_S(\lambda)$. \square

Remark 4.2 Equation (21) and definition (22) reduce to (11) and the subsequent definition of $\ell(\lambda)$ in [18] when the time-dependence of our transition rates is dropped. Conspicuous in the definitions of $M_\lambda(t_0)$ and $\ell(\lambda)$ is that, unlike for L_λ , the adjoint and $\lambda \rightarrow -(1 + \lambda)$ transformations are not inverses, implying that the analogue of H_λ in the proof above—call it \hat{H}_λ —will not satisfy (18). This does not invalidate the fluctuation theorem, however, because we merely require that \hat{H}_λ and $\hat{H}_{-(1+\lambda)}^*$ have the same maximum eigenvalue, not necessarily be equal themselves. The fluctuation theorem for entropy production, in fact, is very straightforward.

Theorem 4.3 *The free energy $c_S^B(\lambda)$ exists, is continuously differentiable $\forall \lambda \in \mathbb{R}$ and satisfies, together with $c_S(\lambda)$, the symmetry relation*

$$c_S(\lambda) = c_S^B(-(1 + \lambda)),$$

implying the fluctuation theorem

$$I_S(z) - I_S^B(-z) = -z, \tag{23}$$

where I_S^B is the continuous and strictly convex good rate function of the entropy production under the backward process.

Proof $P_{[0,t]}$ and $P_{[0,t]}^B$ are mutually absolutely continuous by the ergodic consistency of $A(t)$ [15]. That their Radon-Nikodym derivatives with respect to each other are then reciprocals almost surely is a basic measure-theoretic fact [11]. Together, these justify the moment generating function symmetry

$$\begin{aligned} \mathbb{E}_{\pi,0}(e^{\lambda S(0,t)}) &= \int_{\Omega} \left(\frac{dP_{[0,t]}}{dP_{[0,t]}^B}(\omega) \right)^\lambda P_{[0,t]}(d\omega) = \int_{\Omega} \left(\frac{dP_{[0,t]}}{dP_{[0,t]}^B}(\omega) \right)^\lambda \frac{dP_{[0,t]}}{dP_{[0,t]}^B}(\omega) P_{[0,t]}^B(d\omega) \\ &= \int_{\Omega} \left(\frac{dP_{[0,t]}^B}{dP_{[0,t]}}(\omega) \right)^{-(1+\lambda)} P_{[0,t]}^B(d\omega) = \mathbb{E}_{\mu(\cdot,t),t}^B(e^{-(1+\lambda)S^B(0,t)}), \end{aligned}$$

which is equivalent to the fundamental fluctuation relation of Harris and Schütz [17], cast instead in a quantum Hamiltonian formalism. We then immediately have the free energy symmetry

$$\begin{aligned} c_S(\lambda) &= \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_{\pi,0}(e^{\lambda S(0,t)}) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_{\mu(\cdot,t),t}^B(e^{-(1+\lambda)S^B(0,t)}) \\ &= c_S^B(-(1 + \lambda)), \end{aligned}$$

in which the existence and continuous differentiability of the RHS comes along for free, along with the stated properties of its corresponding rate function. By the Gärtner-Ellis theorem and equalities analogous to (19), this implies the fluctuation theorem symmetry. \square

Informally, the fluctuation theorem symmetry (23) yields the following generalization of the second law of thermodynamics:

$$\frac{P(S(0, t)/t = z)}{P^B(S^B(0, t)/t = -z)} \sim e^{zI}. \tag{24}$$

This has the interpretation that spontaneous positive fluctuations in the time-averaged entropy production of the forward process are exponentially more likely than spontaneous *negative* fluctuations in the *backward* process, and vice-versa. Thus, for processes driven by time-dependent protocols, the distributions of entropy production fluctuations in universes in which time moves forward versus backward are interrelated. Only when the protocol is symmetric within each driving period (i.e., has no temporal orientation) does (24) reclaim its usual irreversibility interpretation that positive entropy production fluctuations in the forward process are exponentially more likely than negative ones.

Taking $\Delta z = [z - \epsilon, z + \epsilon]$, $-\Delta z = [-z - \epsilon, -z + \epsilon]$ and $|\Delta z| = \epsilon$, we formalize (24) as follows.

Corollary 4.4

$$\lim_{|\Delta z| \rightarrow 0} \lim_{t \rightarrow \infty} \frac{1}{t} \log \frac{P(S(0, t)/t \in \Delta z)}{P^B(S^B(0, t)/t \in -\Delta z)} = z.$$

Proof By both the large deviation upper and lower bounds [6], $\lim_{t \rightarrow \infty} \frac{1}{t} \log P(S(0, t)/t \in J) = -\inf_{\zeta \in J} I_S(\zeta)$ for any closed, bounded interval $J \subset \mathbb{R}$, and similarly for $S^B(0, t)$. Therefore, by (23) and continuity of I_S and I_S^B ,

$$\begin{aligned} \lim_{|\Delta z| \rightarrow 0} \lim_{t \rightarrow \infty} \frac{1}{t} \log \frac{P(S(0, t)/t \in \Delta z)}{P^B(S^B(0, t)/t \in -\Delta z)} &= \lim_{|\Delta z| \rightarrow 0} -\inf_{\zeta \in \Delta z} I_S(\zeta) + \inf_{\zeta \in -\Delta z} I_S^B(\zeta) \\ &= -I_S(z) + I_S^B(-z) = z. \end{aligned} \quad \square$$

5 Time-Averaged Entropy Production and Entropy Production Rate

Our final main result is a derivation of the asymptotic time-averaged entropy production and its associated instantaneous entropy production rate. We begin with the following lemma, which is a well-known large deviations result for countable sequences of random variables [6] and requires only slightly more work in the continuous parameter case.

Lemma 5.1 *The time-averaged entropy production $S(0, t)/t$ converges exponentially and $P_{[0, t]}$ -a.s. to $c'_S(0)$.*

Proof By Theorem 4.1, the rate function

$$I_S(z) = \sup_{\lambda \in \mathbb{R}} \{\lambda z - c_S(\lambda)\} \tag{25}$$

is strictly convex. It is immediate from the definition (20) that $c_S(0) = 0$, and so from the inverse transform $c_S(\lambda) = \sup_{z \in \mathbb{R}} \{\lambda z - I_S(z)\}$ we obtain

$$0 = c_S(0) = \sup_{z \in \mathbb{R}} \{-I_S(z)\} = -\inf_{z \in \mathbb{R}} \{I_S(z)\},$$

which, along with strict convexity, implies that $I_S(z)$ attains its minimum and zero uniquely at some z^* . That $z^* = c'_S(0)$ can be seen from the following argument. Recall from the proof of Corollary 3.5 that $I_S(z) = \lambda_z z - c_S(\lambda_z)$, where λ_z is defined implicitly through $c'_S(\lambda_z) = z$. At the same time, since $c_S(0) = 0$, it is clear that (25) is minimized when $\lambda \equiv \lambda_{z^*} = 0$, or $z^* = c'_S(0)$.

Exponential convergence of $S(0, t)/t$ to $c'_S(0)$ is now immediate, because for any $\epsilon > 0$, the large deviation upper bound

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log P_{[0,t]} \left(\left| \frac{S(0, t)}{t} - c'_S(0) \right| \geq \epsilon/2 \right) \leq - \inf_{|z - c'_S(0)| \geq \epsilon/2} I_S(z) \equiv -\alpha_\epsilon < 0$$

guarantees that for large t ,

$$P_{[0,t]} \left(\left| \frac{S(0, t)}{t} - c'_S(0) \right| \geq \epsilon/2 \right) \leq e^{-\frac{\alpha_\epsilon}{2} t}. \tag{26}$$

(The factor 1/2 multiplying ϵ will be needed in a triangle inequality argument to follow.)

Almost sure convergence now follows from several applications of the Borel Cantelli lemma, which says that if the probabilities of a countable sequence of events are summable, then the events happen finitely often almost surely (i.e., only a finite number of them occur). In this case, we have that the events $|S(0, n)/n - c'_S(0)| \geq \epsilon/2$ occur finitely often, where time has been restricted to the positive integers. To show that the same holds for general t , let $(t_j)_{j \geq 1}$ be any sequence of real numbers tending to ∞ . Then

$$\begin{aligned} &P_{[0,t]} \left(\left| \frac{S(0, t_j)}{t_j} - \frac{S(0, \lfloor t_j \rfloor)}{\lfloor t_j \rfloor} \right| \geq \epsilon/2 \right) \\ &\leq P_{[0,t]} \left(\left| \frac{S(0, t_j)}{t_j} - c'_S(0) \right| + \left| c'_S(0) - \frac{S(0, \lfloor t_j \rfloor)}{\lfloor t_j \rfloor} \right| \geq \epsilon/2 \right) \\ &\leq P_{[0,t]} \left(\left| \frac{S(0, t_j)}{t_j} - c'_S(0) \right| \geq \epsilon/4 \right) + P_{[0,t]} \left(\left| \frac{S(0, \lfloor t_j \rfloor)}{\lfloor t_j \rfloor} - c'_S(0) \right| \geq \epsilon/4 \right). \end{aligned}$$

But by an inequality analogous to (26), these latter probabilities can be bounded by a common exponential for large t_j , indicating, by Borel-Cantelli, that $S(0, t_j)/t_j$ and $S(0, \lfloor t_j \rfloor)/\lfloor t_j \rfloor$ differ by more than $\epsilon/2$ only finitely often (almost surely). The triangle inequality then gives us $|S(0, t_j)/t_j - c'_S(0)| < \epsilon$ for large t , precluding, as the t_j 's were arbitrary, the possibility of $|S(0, t)/t - c'_S(0)| \geq \epsilon$ for large t . This is precisely the definition of almost sure convergence to $c'_S(0)$. \square

Theorem 5.2 *The time-averaged entropy production satisfies*

$$0 \leq \lim_{t \rightarrow \infty} \frac{S(0, t)}{t} = \frac{1}{T} \int_0^T e_p(s) ds < \infty, \tag{27}$$

where the convergence is exponential and $P_{[0,t]}$ -a.s., the instantaneous entropy production rate

$$e_p(s) = \frac{1}{2} \sum_{i,j=1}^N (v(i, s)k_{ij}(s) - v(j, s)k_{ji}(s)) \log \left(\frac{v(i, s)k_{ij}(s)}{v(j, s)k_{ji}(s)} \right) \tag{28}$$

is defined for $0 \leq s \leq t$, and $v(\cdot, t) = \lim_{t \rightarrow \infty} \mu(\cdot, t)$ is T -periodic.

Proof By the lemma, we already have convergence in both senses to $c'_S(0)$, so what remains is to evaluate this derivative. In doing so, we make use of the fact that the free energy is convex and that limits and derivatives commute for sequences of convex functions [6].

$$\begin{aligned} c'_S(0) &= \frac{\partial}{\partial \lambda} \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_{\pi,0}(e^{\lambda S(0,t)}) \Big|_{\lambda=0} = \lim_{t \rightarrow \infty} \frac{1}{t} \frac{\partial}{\partial \lambda} \log \mathbb{E}_{\pi,0}(e^{\lambda S(0,t)}) \Big|_{\lambda=0} \\ &= \lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E}_{\pi,0} S(0,t) = \lim_{t \rightarrow \infty} \frac{1}{t} \mathcal{H}(P_{[0,t]}, P_{[0,t]}^B). \end{aligned}$$

Recall that $\mathcal{H}(P_{[0,t]}, P_{[0,t]}^B)$ here is the relative entropy of $P_{[0,t]}$ with respect to $P_{[0,t]}^B$, which are mutually absolutely continuous by the ergodic consistency criterion on A . By nonnegativity of the relative entropy, we have the left inequality of (27). Our goal is now to factor the Radon-Nikodym derivative $dP_{[0,t]}/dP_{[0,t]}^B$ for finite t . To accomplish this, we note that on the event that the Markov chain jumps m times between times t_0 and t , its path ω is characterized by the states $(\sigma_i(\omega))_{i=0}^m$ it visits in sequence and the waiting times $(\tau_{i+1}(\omega) - \tau_i(\omega))_{i=0}^m$ between them. (Here we have set $\tau_0(\omega) = t_0$ and $\tau_{m+1}(\omega) = t$.) Because the former are discrete and the latter have a density, as well as the fact that the number of jumps $J(\omega) = J([t_0, t])(\omega)$ over the interval is finite a.s., the path measure $P_{[t_0,t]}$ has density

$$f_{[t_0,t]}(\omega) = f_{[t_0,t],J(\omega)}(\sigma_0(\omega), \dots, \sigma_{J(\omega)}(\omega), \tau_1(\omega), \dots, \tau_{J(\omega)}(\omega)) \tag{29}$$

where

$$\begin{aligned} &f_{[t_0,t],m}(\sigma_0, \dots, \sigma_m, \tau_1, \dots, \tau_m) \\ &= P(J([t_0, t]) = m) \mu(\sigma_0, t_0) \times \prod_{i=0}^{m-1} \left(K_{\sigma_i}(\tau_{i+1}) \exp \left[- \int_{\tau_i}^{\tau_{i+1}} K_{\sigma_i}(s) ds \right] \frac{k_{\sigma_i, \sigma_{i+1}}(\tau_{i+1})}{K_{\sigma_i}(\tau_{i+1})} \right) \\ &\quad \times \exp \left[- \int_{\tau_m}^t K_{\sigma_m}(s) ds \right]. \end{aligned} \tag{30}$$

A similar density $f_{[t_0,t]}^-(\omega)$ exists for the measure $P_{[t_0,t]}^-$. For notational clarity in what follows, let us write the restriction of a path ω to an interval $E \subset \mathbb{R}$ as ω_E . Given an $M \in \mathbb{Z}$ and $t > 0$, define the partition variables $s_\ell \equiv s_\ell^{M,t} = \frac{\ell t}{M}$ and scale parameter $h = \frac{t}{M}$. Recalling that $P_{[0,t]}^B(\omega) = P_{[0,t]}^-(r(\omega))$, we have by the Markov property

$$\begin{aligned} &\frac{1}{t} \mathcal{H}(P_{[0,t]}, P_{[0,t]}^B) \\ &= \frac{1}{t} \int_{\Omega} \log \frac{dP_{[0,t]}(\omega)}{dP_{[0,t]}^B(\omega)} P_{[0,t]}(d\omega) \\ &= \frac{1}{t} \int_{\Omega} \log \left(\frac{\mu(\omega_0, 0)}{\mu^-(r(\omega)_0, 0)} \prod_{\ell=0}^{M-1} \frac{f_{(s_\ell, s_{\ell+1}]}(\omega_{(s_\ell, s_{\ell+1}]}) | X_{s_\ell} = \omega_{s_\ell})}{f_{(s_\ell, s_{\ell+1}]}^-(r(\omega)_{(s_\ell, s_{\ell+1}]}) | X_{s_\ell} = r(\omega)_{s_\ell})} \right) P_{[0,t]}(d\omega) \\ &= \frac{1}{t} \sum_{\ell=0}^{M-1} \int_{\Omega} \log \left(\frac{\mu(\omega_{s_\ell}, s_\ell) f_{(s_\ell, s_{\ell+1}]}(\omega_{(s_\ell, s_{\ell+1}]}) | X_{s_\ell} = \omega_{s_\ell})}{\mu(\omega_{s_{\ell+1}}, s_{\ell+1}) f_{(s_{M-1-\ell}, s_{M-\ell}]}^-(r(\omega)_{(s_{M-1-\ell}, s_{M-\ell}]} | X_{s_{M-1-\ell}} = r(\omega)_{s_{M-1-\ell}})} \right) \\ &\quad \times P_{[0,t]}(d\omega). \end{aligned}$$

Here we have rewritten the boundary term $\frac{\mu(\omega_0,0)}{\mu^{-}(r(\omega)_{0,0})} = \frac{\mu(\omega_0,0)}{\mu(\omega_r,t)}$ as the telescoping product $\prod_{\ell=0}^{M-1} \frac{\mu(\omega_{s_\ell},s_\ell)}{\mu(\omega_{s_{\ell+1}},s_{\ell+1})}$ and reversed the order of the density terms in the denominator. On the sets

$$\Omega_{\ell,t,M,i,j} = \{\omega \in \Omega : \omega_{s_\ell} = \omega_{s_{\ell-}} = i \text{ and } \omega_{s_{\ell+1}} = \omega_{s_{\ell+1-}} = j\},$$

which partition Ω up to a set of zero measure (the s_ℓ are continuity points of X_t a.s.), the densities on top and bottom and the measure integrated against simplify so that we obtain

$$\begin{aligned} & \frac{1}{t} \sum_{\ell=0}^{M-1} \sum_{i,j=1}^N \int_{\Omega_{\ell,t,M,i,j}} \log \frac{\mu(i,s_\ell)[P(X_{s_{\ell+1}}=j|X_{s_\ell}=i) + o(h)]}{\mu(j,s_{\ell+1})[P^-(X_{s_{M-\ell}}=i|X_{s_{M-1-\ell}}=j) + o(h)]} P_{[0,t]}(d\omega) \\ &= \frac{1}{t} \sum_{\ell=0}^{M-1} \sum_{i,j=1}^N \log \frac{\mu(i,s_\ell)k_{ij}(s_\ell)h + o(h)}{\mu(j,s_{\ell+1})k_{ji}^-(s_{M-1-\ell})h + o(h)} P_{[0,t]}(\Omega_{\ell,t,M,i,j}) \\ &= \frac{1}{M} \sum_{\ell=0}^{M-1} \frac{1}{h} \sum_{i,j=1}^N \log \frac{\mu(i,s_\ell)k_{ij}(s_\ell)}{\mu(j,s_{\ell+1})k_{ji}^-(s_{M-1-\ell})} (\mu(i,s_\ell)k_{ij}(s_\ell)h + o(h)) \\ &= \frac{1}{t} \sum_{\ell=0}^{M-1} h \sum_{i,j=1}^N \mu(i,s_\ell)k_{ij}(s_\ell) \log \frac{\mu(i,s_\ell)k_{ij}(s_\ell)}{\mu(j,s_{\ell+1})k_{ji}(s_{\ell+1})}. \end{aligned}$$

The $o(h)$ correction in the first line is justified by taking $t \downarrow t_0$ in (30) and noting that $P(J([t_0, t_0 + h]) = m) = o(h^m)$. Letting $M \uparrow \infty$ now in the Riemann sum, by continuity of $\mu(\cdot, t)$ and the rates,

$$\frac{1}{t} \mathcal{H}(P_{[0,t]}, P_{[0,t]}^B) = \frac{1}{t} \int_0^t \sum_{i,j=1}^N \mu(i,s)k_{ij}(s) \log \frac{\mu(i,s)k_{ij}(s)}{\mu(j,s)k_{ji}(s)} ds. \tag{31}$$

By Remark 3.4, for all initial distributions π , $\mu(\cdot, t)$ is asymptotic to a periodic distribution $\nu(\cdot, t)$ equal to the RHS of (13), with which we may replace it in the expression above in the limit $t \rightarrow \infty$. Let us now take $\phi(s)$ to be the integrand on the RHS of (31), so that $\phi(s) \sim e_p(s)$. Given $\epsilon > 0$ and choosing an M large enough so that $|\phi(s) - e_p(s)| < \frac{\epsilon}{2}$ for $s > M$, for sufficiently large t

$$\begin{aligned} \frac{1}{t} \int_0^t |\phi(s) - e_p(s)| ds &\leq \frac{1}{t} \int_0^M |\phi(s) - e_p(s)| ds + \frac{1}{t} \int_M^t \frac{\epsilon}{2} ds \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \left(\frac{t-M}{t} \right) < \epsilon. \end{aligned}$$

We may then finally conclude, by periodicity of $e_p(s)$, that

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{S(t)}{t} &= \lim_{t \rightarrow \infty} \frac{1}{t} \mathcal{H}(P_{[0,t]}, P_{[0,t]}^B) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \phi(s) ds \\ &= \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t e_p(s) ds = \frac{1}{T} \int_0^T e_p(s) ds. \end{aligned} \quad \square$$

Remark 5.3 The result (27) that we obtain above is the continuous time analogue of Ge et al.'s complete entropy production rate [16] for discrete time, periodically time-inhomogeneous Markov chains. The form of $e_p(s)$ itself is a time-dependent generalization of both (33) in [14], where $e_p(s)$ is interpreted as the difference between backward and forward dynamical entropy rates, and (4) in [18]. From a dynamical systems perspective, $\nu(\cdot, t)$ is an attracting limit cycle for $\mu(\cdot, t)$ in the space of distributions on $\{1, \dots, N\}$, representing a sort of periodic steady state. The mean entropy produced along this cycle is the mean entropy produced by $\mu(\cdot, t)$ in the long time limit. We see below that when the time dependence of the transition rates is dropped, the limit cycle collapses to a fixed point and $e_p(s)$ reduces to the constant entropy production rate of homogeneous chains.

Corollary 5.4 *When $k_{ij}(t) \equiv k_{ij}, \forall 0 \leq i, j \leq N$,*

$$0 \leq \lim_{t \rightarrow \infty} \frac{S(0, t)}{t} = e_p \equiv \frac{1}{2} \sum_{i, j=1}^N (\mu(i)k_{ij} - \mu(j)k_{ji}) \log \left(\frac{\mu(i)k_{ij}}{\mu(j)k_{ji}} \right), \tag{32}$$

where the convergence is exponential and $P_{[0,1]}$ -almost sure and μ is the unique invariant distribution of X_t .

Proof The existence and uniqueness of μ are guaranteed by the ergodic theorem. We then have $\nu(\cdot, t) = \lim_{t \rightarrow \infty} \mu(\cdot, t) = \mu$, from which the result immediately follows by Theorem 5.2. □

6 Beyond Periodic Driving

The key modeling assumption made in this paper is that the transition rates driving the Markov chain are time-periodic. Despite the wealth of interesting physical and biological processes characterized by this type of driving, examples of which were given in the introduction, an open question remains as to whether there exists a reasonable set of minimal conditions on the transition rates that are weaker than periodicity but still guarantee existence of an AFT. By “reasonable”, we mean that the conditions be both concise and easily verifiable for a wide range of applications. It is well known that transient fluctuation theorems make no assumptions about the protocol driving the process, suggesting that the minimal conditions necessary for an AFT to hold truly would be minimal. An initial guess might be that uniform continuity and boundedness are sufficient, which would prevent the number of jumps of the chain from growing faster than linearly in time as well as preventing pathological behavior due to discontinuous or infinitely rapid driving. The following result shows that this guess is incorrect.

Proposition 6.1 *Given arbitrary constants $0 < \alpha < \beta$, there exists a Markov chain whose transition rates are uniformly continuous and bounded between α and β , but whose time-averaged entropy production does not satisfy an AFT.*

Proof The plan is to take $0 < \alpha < \beta$ as given and then construct a chain with the properties above. We begin by choosing distinct constants $\alpha < k_{ij} < \beta$ for $1 \leq i, j \leq N$ such that the generator $A_c = (k_{ij})$ does not satisfy detailed balance, which of course is always possible. Note that A_c is irreducible because all its entries are positive, and so it possesses a unique ergodic distribution. We now choose a number $\gamma > 1$ such that $\alpha < \gamma k_{ij} < \beta$ for all i and j ,

and let τ be the mixing time of the chain generated by γA_c (the typical time until convergence to its ergodic distribution), which is equal to the reciprocal of the second largest real part of the eigenvalues of γA_c . The generator $A(t)$ of our chain is constructed as follows. With $t_0 \equiv 0$, we choose $t_1 \gg \tau$ and then iteratively set $t_k = kt_{k-1}$. $A(t)$ is initially defined to be A_c for $t_{2k} \leq t < t_{2k+1}$ and γA_c for $t_{2k+1} \leq t < t_{2k}$. We then modify it by smoothing out the discontinuities in any manner that leaves $A(t)$ uniformly continuous in time.

In order to prove that $S(0, t)/t$ does not satisfy a large deviation property and, hence, an AFT, we show that its free energy $c_S(\lambda)$ fails to converge almost everywhere. To begin, define $c_S(\lambda, t) = \frac{1}{t} \log \mathbb{E}_{\pi,0}(e^{\lambda S(0,t)})$ so that $c_S(\lambda) = \lim_{t \rightarrow \infty} c_S(\lambda, t)$. Now let e_p denote the instantaneous entropy production rate of a process generated by A_c , which is defined by (32) and must be positive because the rates k_{ij} do not satisfy detailed balance. Note that e_p is homogeneous of degree 1 with respect to the rates, so that the entropy production rate of the process generated by γA_c is γe_p . Having defined the t_k to grow such that $t_{k-1}/t_k \rightarrow 0$, over timescales much longer than the mixing times for A_c and γA_c , we see by Corollary 5.4 that

$$S(0, t) = \begin{cases} te_p + o(1), & t_{2k} \leq t < t_{2k+1} \\ t\gamma e_p + o(1), & t_{2k+1} \leq t < t_{2k}, \end{cases}$$

where $o(1)$ denotes a term that vanishes as $t \rightarrow \infty$. This implies that for $\lambda > 0$,

$$\begin{aligned} \liminf_{t \rightarrow \infty} c_S(\lambda, t) &= \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_{\pi,0}(e^{\lambda(te_p + o(1))}) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_{\pi,0}(e^{\lambda te_p}) \\ &< \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_{\pi,0}(e^{\lambda(t\gamma e_p + o(1))}) = \limsup_{t \rightarrow \infty} c_S(\lambda, t) \end{aligned}$$

and hence $c(\lambda)$ does not exist. The inequality is reversed for $\lambda < 0$, so we see that the free energy only exists for the trivial value $\lambda = 0$. □

Since uniform continuity and boundedness are implied by continuity and periodicity, the preceding result shows that the minimal conditions on the transition rates are closer to those assumed in this paper than one might initially suspect. What periodicity guarantees but uniform continuity and boundedness do not is that the rates cannot be tuned over arbitrarily long timescales, exactly the loophole we exploited above. It remains to be seen, however, how this condition can be formulated more precisely and whether there are other conditions must be included in our minimal set.

Acknowledgements The authors would like to thank Kenneth Lange and Hua Zhou for helpful discussions. This work was supported by grants from the NSF (DMS-0349195) and the NIH (K25 AI41935), as well as the VIGRE Graduate Fellowship.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Astumian, R.D.: Thermodynamics and kinetics of a Brownian motor. *Science* **276**, 917 (1997)
2. Bonetto, F., Gallavotti, G., Giuliani, A., Zamponi, F.: Chaotic hypothesis, fluctuation theorem and singularities. *J. Stat. Phys.* **123**, 39 (2006)

3. Chen, B., Shen, X., Li, Y., Sun, L., Yin, Z.: Dynamic theory for the mesoscopic electric circuit. *Phys. Lett. A* **335**, 103 (2005)
4. Crooks, G.E.: Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E* **60**, 2721 (1999)
5. Ellis, R.S.: Large deviations for a general class of random vectors. *Ann. Probab.* **12**, 1 (1984)
6. Ellis, R.S.: *Entropy, Large Deviations, and Statistical Mechanics*. Springer, New York (1986)
7. Ellis, R.S., Haven, K., Turkington, B.: Large deviation principles and complete equivalence and non-equivalence results for pure and mixed ensembles. *J. Stat. Phys.* **101**, 999 (2000)
8. Evans, D.J., Searles, D.J.: The fluctuation theorem. *Adv. Phys.* **51**, 1529 (2002)
9. Evans, D.J., D Cohen, E.G., Morris, G.P.: Probability of second law violations in steady flows. *Phys. Rev. Lett.* **71**, 2401 (1993)
10. Feng, E.H., Crooks, G.E.: Length of time's arrow. *Phys. Rev. Lett.* **101**, 090602 (2008)
11. Folland, G.B.: *Real Analysis*. Wiley, New York (1999)
12. Gallavotti, G., Cohen, E.G.D.: Dynamical ensembles in stationary states. *J. Stat. Phys.* **80**, 931 (1995)
13. Gammaitoni, L., Hänggi, P., Jung, P., Marchesoni, F.: Stochastic resonance. *Rev. Mod. Phys.* **70**, 223 (1998)
14. Gaspard, P.: Time-reversed dynamical entropy and irreversibility in Markovian random processes. *J. Stat. Phys.* **117**, 599 (2004)
15. Ge, H., Jiang, D.: The transient fluctuation theorem of sample entropy production for general stochastic processes. *J. Phys. A* **40**, F713 (2007)
16. Ge, H., Jiang, D., Qian, M.: A simple discrete model of Brownian motors: time-periodic Markov chains. *J. Stat. Phys.* **123**, 831 (2006)
17. Harris, R.J., Schütz, G.M.: Fluctuation theorems for stochastic dynamics. *J. Stat. Mech.* P07020 (2007)
18. Jiang, D., Qian, M., Zhang, F.: Entropy production fluctuations of finite Markov chains. *J. Math. Phys.* **4**, 4176 (2003)
19. Jiang, D.-Q., Qian, M., Qian, M.-P.: *Mathematical Theory of Nonequilibrium Steady States*. Springer, Berlin (2004)
20. Joubaud, S., Garnier, N.B., Ciliberto, S.: Fluctuation theorems for harmonic oscillators. *J. Stat. Mech.* P09018 (2007)
21. Kurchan, J.: Fluctuation theorem for stochastic dynamics. *J. Phys. A* **31**, 3719 (1998)
22. Kurchan, J.: A quantum fluctuation theorem. [arXiv:cond-mat/0007360v2](https://arxiv.org/abs/cond-mat/0007360v2) [cond-mat.stat-mech]
23. Lax, P.D.: *Linear Algebra*. Wiley, New York (1997)
24. Lebowitz, J.L., Spohn, H.: A Gallavotti-Cohen type symmetry in the large deviation functional for stochastic dynamics. *J. Stat. Phys.* **95**, 333 (1999)
25. Maes, C.: The fluctuation theorem as a Gibbs property. *J. Stat. Phys.* **95**, 367 (1999)
26. Maes, C., Netočný, K.: Time-reversal and entropy. *J. Stat. Phys.* **110**, 269 (2003)
27. Maes, C., Redig, F., Van Moffaert, A.: On the definition of entropy production, via examples. *J. Math. Phys.* **41**, 1528 (2000)
28. Price, H.: *Time's Arrow and Archimedes' Point*. Oxford University Press, Oxford (1996)
29. Puglisi, A., Rondoni, L., Vulpiani, A.: Relevance of initial and final conditions for the fluctuation relation in Markov processes. *J. Stat. Mech.* P08010 (2006)
30. Rákos, A., Harris, R.J.: On the range of validity of the fluctuation theorem for stochastic Markovian dynamics. *J. Stat. Mech.* P05005 (2008)
31. Renshaw, E.: *Modelling Biological Populations in Space and Time*. Cambridge University Press, Cambridge (1991)
32. Schuler, S., Speck, T., Tietz, C., Wrachtrup, J., Seifert, U.: Experimental test of the fluctuation theorem for a driven two-level system with time-dependent rates. *Phys. Rev. Lett.* **94**, 180602 (2005)
33. Seifert, U.: Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Phys. Rev. Lett.* **95**, 040602 (2005)
34. Sevick, E.M., Prabhakar, R., Stephen, R., Debra, W., Searles, J.: Fluctuation theorems. *Annu. Rev. Phys. Chem.* **59**, 603 (2008)
35. Singh, N.: Onsager-Machlup theory and work fluctuation theorem for a harmonically driven Brownian particle. *J. Stat. Phys.* **131**, 405 (2008)
36. Serman, G.: *An Introduction to Quantum Field Theory*. Cambridge University Press, Cambridge (1993)
37. Tietz, C., Schuler, S., Seifert, U., Wrachtrup, J.: Measurement of stochastic entropy production. *Phys. Rev. Lett.* **97**, 050602 (2006)
38. Touchette, H.: The large deviation approach to statistical mechanics. [arXiv:0804.0327v1](https://arxiv.org/abs/0804.0327v1) [cond-mat.stat-mech]
39. van Zon, R., Cohen, E.G.D.: Extension of the fluctuation theorem. *Phys. Rev. Lett.* **91**, 110601 (2003)
40. van Zon, R., Cohen, E.G.D.: Extended heat-fluctuation theorems for a system with deterministic and stochastic forces. *Phys. Rev. E* **69**, 056121 (2004)
41. Verhulst, F.: *Nonlinear Differential Equations and Dynamical Systems*. Springer, New York (1996)