Ann. Henri Poincaré 19 (2018), 3197–3238

1424-0637/18/103197-42
published online August 3, 2018
https://doi.org/10.1007/s00023-018-0705-3

Annales Henri Poincaré

Level 2.5 Large Deviations for Continuous-Time Markov Chains with Time Periodic Rates

Lorenzo Bertini, Raphael Chetrite, Alessandra Faggionato and Davide Gabrielli

Abstract. We consider an irreducible continuous-time Markov chain on a finite state space and with time periodic jump rates and prove the joint large deviation principle for the empirical measure and flow and the joint large deviation principle for the empirical measure and current. By contraction, we get the large deviation principle of three types of entropy production flow. We derive some Gallavotti–Cohen duality relations and discuss some applications.

1. Introduction

Periodically driven Markov processes are a common setup for several small systems, such as artificial molecular motors. Unlike their biological counterparts, artificial molecular motors are non-autonomous and work under the effect of time periodic externally modulated stimuli such as temperature, laser light, chemical environment. Periodic forcing is fundamental also in relation to micro-sized engines. For example, experimental heat engines driven by periodic temperature variations have been realized in [7,49] and experimental molecular pumps with periodic modulation of hamiltonian have been realized in [23,43]. Due to their numerous applications inside nanotechnologies, in the last years periodically driven Markov processes have received much attention in the thermodynamic theory of small systems or, equivalently, stochastic thermodynamics [65,66]. Several theoretical results have been obtained for what concerns linear response and Onsager reciprocity relations [8,26,37,56,57,59,71,72], and no-go theorems in stochastic pumping [11,37,45,58,62]. Time periodic forcing

The work of L. Bertini and A. Faggionato has been supported by PRIN 20155PAWZB “Large Scale Random Structures,” while the work of R. Chetrite has been supported by the French Ministry of Education Grant ANR-15-CE40-0020-01.

is also at the basis of stochastic resonance phenomena [28], i.e., of the amplification of a weak periodic signal by means of noise.

We consider here an irreducible continuous-time Markov chain on a finite state space V with time periodic jump rates, having period T_0. We focus on large deviations (shortly, LD) at large times of the empirical measure, flow and current. Roughly, referring to a time window $[0, T]$, for each $y \in V$ the empirical measure $\bar{\mu}_T(y)$ counts the fraction of time spent by the system in the state y. For each pair of states $y, z \in V$ the empirical flow $\bar{Q}_T(y, z)$ counts the number of jumps from y to z per unit time, while the empirical current $\bar{J}_T(y, z)$ is given by the difference $\bar{Q}_T(y, z) - \bar{Q}_T(z, y)$. The above objects enter naturally in several applications. Considering, for example, the dynamics of a molecular motor for which an ATP hydrolysis takes place simultaneously to a transitions from state y to state z, $TQ_T(y, z)$ gives the number of hydrolyzed ATP's in the time window $[0, T]$. To have explicit LD functionals, we consider also extended versions $\mu^{(n)}, Q^{(n)}, J^{(n)}$ of the above empirical measure, flow and current, keeping record of time t apart integer multiples of the period T_0. More precisely, $\mu^{(n)}(y, \mathrm{d}t)$ is defined as the time per period spent at state y during the n infinitesimal time intervals $[t, t + \mathrm{d}t)$, $[t + T_0, t + T_0 + \mathrm{d}t)$,..., $[t + (n-1)T_0, t + (n-1)T_0 + \mathrm{d}t)$. $Q^{(n)}(y, z, \mathrm{d}t)$ equals the number, per period, of jumps from y to z performed in the above n infinitesimal time intervals, while $J^{(n)}(y, z, \mathrm{d}t)$ is defined as the difference $Q^{(n)}(y, z, \mathrm{d}t) - Q^{(n)}(z, y, \mathrm{d}t)$.

Although our initial objects of investigation are given by the empirical measure, flow and current, to get control on large deviations it is crucial to include more information and deal with their extended versions. Indeed, the technical core of our analysis is the derivation of the LD principle for the joint extended empirical measure and flow $\left(\mu^{(n)}, Q^{(n)}\right)$ as n goes to ∞ (cf. Theorem 2). Roughly, we get that

$$\mathbb{P}\left((\mu^{(n)}, Q^{(n)}) \approx (\mu, Q)\right) \asymp e^{-nI(\mu,Q)}, \qquad n \gg 1,$$

for a suitable explicit rate functional $I(\cdot, \cdot)$. By contraction, we obtain the LD principle for the joint extended empirical measure and current $\left(\mu^{(n)}, J^{(n)}\right)$ (cf. Theorem 3). These LD principles correspond to level 2.5 (see below) and remarkably admit explicit LD rate functionals. By contraction, we derive LD principles for the empirical measure, flow and current $\bar{\mu}_T$, $\bar{Q}_T$, $\bar{J}_T$ (cf. Theorem 1 , Remark 3.2 and Remark 3.10). As an application of the above results, after introducing several forms of entropy production, we obtain the associated LD principles by contraction. Moreover, we derive some Gallavotti–Cohen duality relations at the level 2.5 and show that, by projection, these relations imply other Gallavotti–Cohen duality relations for the entropy production rate and some of them are new (cf. Theorem 4, Corollaries 5.2 and 5.3). In addition, we discuss in detail the case of a 2-state model. We point out that for a time periodic symmetric protocol, Gallavotti–Cohen duality relations for the entropy production rate have been experimentally verified in [64] and theoretically analyzed in stochastic models in [69,70]. Recently, in [59] a Gallavotti–Cohen duality relation for a not symmetric protocol has been obtained. Finally, we point out that [68] provides a first theoretical analysis of level 2.5 large deviations in periodically driven diffusion processes.

Our results are a natural development of the analysis of level 2.5 large deviations for time homogeneous ergodic Markov processes. To explain this issue, below we recall some fundamental results. (In particular, below we refer to time homogeneous processes.) The celebrated papers by Donsker and Varadhan [19–22] have provided a crucial contribution to the large deviation theory for ergodic Markov processes. One is typically interested on the long time behavior of the process, and three possible levels on which the large deviations can be investigated have been identified: level 1 that concerns the fluctuations of additive observables; level 2 that concerns the fluctuations of the empirical measure; level 3 that concerns the fluctuations of the empirical process. These levels have a hierarchical structure, and the large deviations on a lower level can be deduced by projection. As the name implies, level 2.5 lies in between level 2 and level 3 and concerns the joint fluctuations of the empirical measure and empirical flow (or the joint fluctuations of the empirical measure and empirical current). In the simple context of homogeneous continuous-time Markov chains, the empirical flow counts the numbers of jumps between pairs of states. We emphasize that in this case the rate functional for level 1 cannot be expressed in closed form, for level 2 this is possible only in the reversible case, while for level 3 the rate functional is given by the specific relative entropy with respect to the stationary process that gives an explicit but somehow abstract formula. On the other hand, for level 2.5 there is a simple explicit formula that covers both the reversible and non-reversible case, so level 2.5 represents the lowest level admitting an explicit rate functional.

A relevant motivation for the analysis of large deviations at level 2.5 comes from non-equilibrium statistical physics. Indeed, in this context the current flowing through the system is a key observable and exhibits rich and peculiar large deviation behavior, see, e.g., [3,42]. Moreover, the statistics of the entropy production and the Gallavotti–Cohen symmetry cannot be described only in terms of the empirical measure but require also the current [12,46,47]. From a purely probabilistic viewpoint, the level 2.5 has been firstly investigated in [39] in the case of a two-state chain. For a countable state space, the level 2.5 weak large deviation principle has been established in [24]. In the same setting, the large deviation principle is proven in [5] (and further analyzed in [4,6]), while the analogous result for diffusion processes is obtained in [40]. A more general setting with time-dependent empirical measure and flow is considered in [38,61]. We also point out that recently some thermodynamic uncertainty relation [1] and some related universal bound on current fluctuations [54,55] have been derived in [29,30] by using the level 2.5 large deviation principle. Finally, we refer to [4,13] for a further discussion about level 2.5 for time homogeneous Markov processes.

Outline of the Paper In Sect. 2, we describe our main assumptions on the continuous-time Markov chains with time periodic rates. In Sect. 3, we introduce the empirical measure, flow, current and state our main large deviation principles (cf. Theorems 1, 2 and 3). In Sect. 4, we discuss three forms of entropy production, and in Sect. 5, we state the associated Gallavotti–Cohen duality relations (cf. Theorem 4, Corollaries 5.2 and 5.3). In Sect. 6, we apply our general results to the case of two-state continuous-time Markov chains with time periodic rates. The rest of the paper is devoted to the proof of our results. In particular, in Sect. 7 we collect some preliminary facts. In Sect. 8, we prove the upper bound for the LDP stated in Theorem 2 (cf. Eq. (3.13)), the convexity and the lower-semicontinuity of the LD rate functional of Theorem 2, while in Sect. 9 we prove the lower bound (cf. Eq. (3.13)) and the goodness of the same rate functional. Theorem 1 follows easily from Theorem 2 by contraction, and therefore, the proof is omitted. The proofs of Theorems 3 and Theorem 4 are given in Sects. 10 and 11, respectively.

2. Continuous-Time Markov Chains with Time Periodic Rates

We consider a continuous-time Markov chain $\xi = (\xi_t)_{t\in\mathbb{R}_+}$ on a finite state space V, with time periodic jump rates. We call $r(y, z; t)$ the jump rate from y to z at time t and we assume that $r(\cdot, \cdot; t) = r(\cdot, \cdot; t + T_0)$ for some $T_0 > 0$. To have a well-defined process, we assume that $r(y, z; \cdot)$ is a measurable, locally integrable nonnegative function (see below). We also convey that $r(x, x; t) \equiv 0$.

Roughly, the dynamics is defined as follows. Starting from a state x, the Markov chain spends at x a random time τ_1 such that

$$\mathbb{P}(\tau_1 > t) = \exp\left\{-\int_0^t r(x; s)\mathrm{d}s\right\},$$

where

$$r(x; s) := \sum_z r(x, z; s).$$

Knowing that $\tau_1 = t_1$, at time t_1 the Markov chain jumps to a new state x_1 chosen randomly with probability $r(x, x_1; t_1)/r(x; t_1)$; afterward, it waits in x_1 a random time τ_2 such that

$$\mathbb{P}(\tau_2 > t) = \exp\left\{-\int_{t_1}^{t_1+t} r(x_1; s)\mathrm{d}s\right\}.$$

Knowing that $\tau_2 = t_2$, at time t_2 the Markov chain jumps to a new state x_2 chosen randomly with probability $r(x_1, x_2; t_2)/r(x_1; t_2)$, and so on.

Above we have not used the periodicity of the jump rates, and indeed the construction is common to all time inhomogeneous Markov chains. Formally, a time inhomogeneous Markov chain can be seen as a piecewise-deterministic Markov process and its precise definition follows from the general construction in [17]. Indeed, we can introduce the continuous variable $s \in [0, +\infty)$ and describe the state of the system at time t by (ξ_t, s_t) where $s_t := t$. Then, the evolution in $V \times \mathbb{R}_+$ is described by a time homogeneous piecewise-deterministic Markov process with formal generator L

$$Lf(x,s) = \partial_s f(x,s) + \sum_y r(x,y;s)\big[f(y,s) - f(x,s)\big]. \tag{2.1}$$

Following [17], to have a well-defined operator one needs that the jump rates $r(x,y;\cdot)$ are measurable, locally integrable nonnegative functions. Due to [17] time inhomogeneous Markov chains enjoy the strong Markov property.

We denote by E the set of pairs (y,z) such that $r(y,z;t)>0$ for all $t>0$, $y\neq z$. We think of (V,E) as a directed graph. Moreover, we write $\mathcal{S}_{T_0}$ for the set $\mathbb{R}/T_0\mathbb{Z}$, i.e., for the set $[0,T_0]$ with periodic boundary conditions (0 and T_0 have to be identified).

Assumptions *Our assumptions are the following:*

(A1) *If* $r(y,z;t)>0$ *for some* $t>0$*, then* $r(y,z;t)>0$ *for all* $t>0$*;*
(A2) *The directed graph* (V,E) *is strongly connected;*
(A3) *The jump rates are nonnegative measurable functions such that*

$$\max_{(y,z)\in E}\ \sup_{t\in[0,T_0]} r(y,z;t) < \infty, \tag{2.2}$$

$$\min_{(y,z)\in E}\ \inf_{t\in[0,T_0]} r(y,z;t) > 0. \tag{2.3}$$

(A4) *We assume that the set* $\mathcal{D}$ *has zero Lebesgue measure, where* $\mathcal{D}\subset\mathcal{S}_{T_0}$ *is the set of discontinuity points of the jump rates* $\mathcal{S}_{T_0}\ni t\mapsto r(y,z;t)\in[0,\infty)$*, as* y,z *vary in* V*.*

Assumption (A2) means that, given two distinct sites y,z in V, there is a family of vertexes $x_0,x_1,\dots,x_n$ such that $x_0=y$, $x_n=z$ and $(x_i,x_{i+1})\in E$ for all $i=0,1,\dots,n-1$.

We point out that assumption (A4) is used only to derive Lemma 8.2.

Trivially, the discrete-time process $\widetilde{\xi}=(\widetilde{\xi}_n)_{n\geq 0}$, with $\widetilde{\xi}_n:=\xi_{nT_0}$, is a time homogeneous Markov chain. We write $\widetilde{p}(y,z)$, $y,z\in V$, for its jump probabilities. Since V is finite and (V,E) is strongly connected, $\widetilde{\xi}$ admits a unique invariant distribution π_0, i.e., a unique probability measure π_0 on V such that

$$\sum_{y\in V}\pi_0(x)\,\widetilde{p}(x,y) = \sum_{y\in V}\pi_0(y)\,\widetilde{p}(y,x) \qquad \forall\, x\in V. \tag{2.4}$$

Note that the Markov chain $\widetilde{\xi}$ starting with the invariant distribution π_0 is stationary, i.e., its law is invariant under time shifts (cf. Th.1.7.1 in [52]). As a by-product of this fact, the Markov property fulfilled by ξ and the T_0-periodicity of the jump rates, one easily gets that the Markov chain ξ starting with initial distribution π_0 is T_0-stationary, i.e., its law is invariant under time translations along times $T_0,2T_0,3T_0,\dots$. In particular, when ξ starts with distribution π_0, the law π_t of ξ_t is a T_0-periodic function from $\mathbb{R}_+$ to the space $\mathcal{P}(V)$ of probability measures on V. We point out that π_0 is indeed the only initial distribution for which the Markov chain ξ starting at π_0 is T_0-periodic; hence, we call the associated law of $\xi=(\xi_t)_{t\geq 0}$ on the space of càdlàg paths $D(\mathbb{R}_+;V)$ the *oscillatory steady state* (sometimes, as in [67], this state is called *non-equilibrium oscillatory state*, shortly NOS).

In what follows, we set $\pi := \pi_t dt$. π is a nonnegative measure on $V \times \mathcal{S}_{T_0}$ with total mass T_0 (shortly, $\pi \in \mathcal{M}_{+,T_0}(V \times \mathcal{S}_{T_0})$). Given a probability measure ν on V, we write $\mathbb{P}_\nu$ for the law of the Markov chain $(\xi_t)_{t\geq 0}$ with initial distribution ν, and we simply write $\mathbb{P}_x$ if $\nu = \delta_x$. The associated expectations are denoted by $\mathbb{E}_\nu$ and $\mathbb{E}_x$, respectively.

2.1. Graphical Construction

We conclude by providing a graphical construction of the continuous-time Markov chain $(\xi_t)_{t\in\mathbb{R}_+}$, which will be useful in what follows.

To each $(y,z) \in E$ we associate a Poisson process of rate $\lambda(y,z) = \sup_{t\in[0,T_0]} r(y,z;t)$. We write

$$\mathcal{T}_{y,z} = \{t_{y,z}^{(1)} < t_{y,z}^{(2)} < \cdots\}$$

for the jump times of the above Poisson process. Let us write

$$\mathcal{T}_y = \{t_y^{(1)} < t_y^{(2)} < \cdots\}$$

for the superposition $\cup_z \mathcal{T}_{y,z}$. It is known that $\mathcal{T}_y$ is a Poisson point process on $(0,\infty)$ with rate $\lambda(y) := \sum_z \lambda(y,z)$ (i.e., $\mathcal{T}_y$ is the set of jump times of a Poisson process with rate $\lambda(y)$). Note that $\lambda(y) < \infty$ due to (2.2). For each $(y,z) \in E$ consider also a sequence of i.i.d. random variables $\mathcal{U}_{y,z} = (U_{y,z}^{(k)})_{k\geq 1}$ uniformly distributed on $[0,1]$. The random objects given by $\mathcal{U}_{y,z}$, $\mathcal{T}_{y,z}$ with (y,z) varying in E must be all independent.

Then, the graphical construction is the following. Suppose that $t = 0$ or that the chain has been updated at time t and its state at time t is y. Let s be the minimum of the set $\mathcal{T}_y \cap (t,+\infty)$ and let k, z be such that $s = t_{y,z}^{(k)}$ (they are well defined a.s.). Then, $s = t_{y,z}^{(k)}$ is an update time and the update is the following: if $U_k \leq \frac{r(y,z;s)}{\lambda(y,z)}$, then we let $\xi_s := z$; otherwise, we let $\xi_s := y$. After the update, the algorithm starts afresh.

3. Large Deviation Principles

3.1. Joint LDP for the Empirical Measure and Flow

Definition 3.1. *Given $T > 0$, to each path $X \in D(\mathbb{R}_+; V)$ we associate the empirical measure $\bar\mu_T(X) \in \mathcal{P}(V)$ and the empirical flow $\bar Q_T(X) \in \mathbb{R}_+^E$ defined as*

$$\bar\mu_T(X) = \frac{1}{T}\int_0^T \delta_{X_t}\, dt, \qquad \bar Q_T(X)(y,z) = \frac{1}{T}\sum_{\substack{t\in[0,T]:\\ X_{t-}\neq X_t}} \mathbb{1}\big((X_{t-}, X_t) = (y,z)\big).$$

Let $\Phi\colon \mathbb{R}_+ \times \mathbb{R}_+ \to [0,+\infty]$ be the function defined by

$$\Phi(q,p) := \begin{cases} q\log\dfrac{q}{p} - (q-p) & \text{if } q,p \in (0,+\infty),\\ p & \text{if } q = 0,\ p \in [0,+\infty),\\ +\infty & \text{if } p = 0 \text{ and } q \in (0,+\infty). \end{cases} \tag{3.1}$$

For $p > 0$, $\Phi(\cdot, p)$ is a nonnegative strictly convex function and is zero only at $q = p$. Indeed, since $\Phi(q, p) = \sup_{s\in\mathbb{R}} \{qs - p(e^s - 1)\}$, $\Phi(\cdot, p)$ is the rate functional for the LDP of the sequence N_T/T as $T \to +\infty$, $(N_t)_{t\in\mathbb{R}_+}$ being a Poisson process with parameter p.

Given $t \in [0, T_0)$, let $I_t \colon \mathcal{P}(V) \times \mathbb{R}_+^E \to [0, +\infty]$ be the functional defined by

$$I_t(\bar\mu, \bar Q) := \sum_{(y,z)\in E} \Phi\big(\bar Q(y, z), \bar\mu(y) r(y, z; t)\big). \tag{3.2}$$

Theorem 1. *For each $x \in V$, by taking T of the form $T = nT_0$ with n integer, as $T \to +\infty$ the family of probability measures $\{\mathbb{P}_x \circ (\bar\mu_T, \bar Q_T)^{-1}\}$ on $\mathcal{P}(V) \times \mathbb{R}_+^E$ satisfies a large deviation principle with speed T and good and convex rate functional $\bar I$ defined as*

$$\bar I(\bar\mu, \bar Q) = \inf_{(\mu_t, Q_t)_{t\in\mathcal{S}_{T_0}}} \frac{1}{T_0}\int_0^{T_0} I_t(\mu_t, Q_t)\, dt, \tag{3.3}$$

where the infimum is taken among all measurable functions $\mathcal{S}_{T_0} \ni t \to (\mu_t, Q_t) \in \mathcal{P}(V) \times \mathbb{R}_+^E$ such that

$$\begin{cases} \partial_t \mu_t + \operatorname{div} Q_t = 0, \\ \frac{1}{T_0}\int_0^{T_0} \mu_t\, dt = \bar\mu, \\ \frac{1}{T_0}\int_0^{T_0} Q_t\, dt = \bar Q. \end{cases} \tag{3.4}$$

Theorem 1 follows easily by contraction from Theorem 2 below; hence, we omit the proof.

We give some comments on the notation used in Theorem 1. First, we recall that the infimum of the empty set equals $+\infty$ by definition. We also recall that given $A \in \mathbb{R}_+^E$, the divergence $\operatorname{div} A : V \to \mathbb{R}$ is defined as

$$\operatorname{div} A(y) = \sum_{z:(y,z)\in E} A(y, z) - \sum_{z:(z,y)\in E} A(z, y). \tag{3.5}$$

Below we will often use the convention that given a function $B : E \to \mathbb{R}$, we set $B(y, z) := 0$ if $(y, z) \notin E$. For example, due to this convention, we can rewrite (3.5) as $\operatorname{div} A(y) = \sum_z A(y, z) - \sum_z A(z, y)$. Finally, the above *continuity equation* $\partial_t \mu_t + \operatorname{div} Q_t = 0$ in (3.4) is thought of in weak sense, i.e., using the time T_0-periodicity

$$\int_0^{T_0} \sum_y \mu_s(y) \partial_s f(y, s) \mathrm{d}s = \int_0^{T_0} \sum_y \operatorname{div} Q_s(y) f(y, s) \mathrm{d}s, \tag{3.6}$$

for any C^1 function $f : V \times \mathcal{S}_{T_0} \to \mathbb{R}$. Here and in what follows, the C^1-regularity refers to time.

We finally observe that if $\operatorname{div} \bar Q \neq 0$ then $\bar I(\bar\mu, \bar Q) = +\infty$. Indeed by taking time average of the continuity equation in (3.4) and using that $t \mapsto \mu_t$ is defined on $\mathcal{S}_{T_0}$ (and therefore can be thought of as a T_0-periodic function), we deduce $\operatorname{div} \bar Q = 0$.

Remark 3.2. By contraction, one derives from Theorem 1 both a LDP for the empirical measure $\bar{\mu}_T$ and a LDP for the empirical flow $\bar{Q}_T$ (cf. [4] for the corresponding contraction in the time homogeneous case).

Remark 3.3. By the goodness of the rate function I in Theorem 2, the infimum in (3.3) is achieved whenever (3.4) admits a solution. In particular, by the goodness of I, we have that $\bar{I}(\bar{\mu}, \bar{Q}) = 0$ if and only if there exists a pair $\mu = \mu_t \mathrm{d}t$, $Q = Q_t \mathrm{d}t$ solving (3.4) and such that $I(\mu, Q) = 0$. As a by-product with Remark 3.6 below, we conclude that $\bar{I}(\bar{\mu}, \bar{Q}) = 0$ if and only if $\bar{\mu} = \frac{1}{T_0}\int_0^{T_0} \pi_t \,\mathrm{d}t$ and $\bar{Q}(y,z) = \frac{1}{T_0}\int_0^{T_0} \pi_t(y) r(y,z;t)\,\mathrm{d}t$ for each $(y,z) \in E$.

3.2. Joint LDP for the Extended Empirical Measure and Flow

We introduce the space $\mathcal{M}_{+,T_0}(V \times \mathcal{S}_{T_0})$ as the family of nonnegative measures on $V \times \mathcal{S}_{T_0}$ with total mass equal to T_0, and the space $\mathcal{M}_+(E \times \mathcal{S}_{T_0})$ as the family of nonnegative measures on $E \times \mathcal{S}_{T_0}$ with finite total mass. Both spaces are endowed with the weak topology, i.e., $\nu_n \to \nu$ if and only if $\nu_n(f) \to \nu(f)$ for any bounded continuous function f (by compactness, continuous functions on $V \times \mathcal{S}_{T_0}$ and on $E \times \mathcal{S}_{T_0}$ are automatically bounded). We will often use the trivial identifications $\mathcal{M}_+(V \times \mathcal{S}_{T_0}) \sim \mathcal{M}_+(\mathcal{S}_{T_0})^V$ and $\mathcal{M}_+(E \times \mathcal{S}_{T_0}) \sim \mathcal{M}_+(\mathcal{S}_{T_0})^E$, as in the following definition:

Definition 3.4. *Given a positive integer n, to each path $X \in D(\mathbb{R}_+; V)$ we associate the* extended empirical measure $\mu^{(n)} \in \mathcal{M}_{+,T_0}(V \times \mathcal{S}_{T_0})$ *and the* extended empirical flow $Q^{(n)} \in \mathcal{M}_+(E \times \mathcal{S}_{T_0})$ *defined by*

$$\mu^{(n)}(x, dt) = \mu_t^{(n)}(x)\,dt \quad \textit{where} \quad \mu_t^{(n)}(x) := \frac{1}{n}\sum_{k=0}^{n-1} \delta_{X_{t+kT_0}}(x), \tag{3.7}$$

$$Q^{(n)}(y,z,B) = \frac{1}{n}\sum_{k=0}^{n-1} \sum_{\substack{t \in B + kT_0 : \\ X_{t-} \neq X_t}} \mathbb{1}\big((X_{t-}, X_t) = (y,z)\big), \tag{3.8}$$

where B is a generic Borel subset $B \subset (0, T_0]$. (In the above formulas, we have used the natural parametrization of $\mathcal{S}_{T_0}$ by $(0, T_0]$.)

We can identify functions $f : V \times \mathcal{S}_{T_0} \to \mathbb{R}$ with functions $f : V \times \mathbb{R}_+ \to \mathbb{R}$ which are T_0-periodic in the time variable. In what follows, when we say that $f : V \times \mathbb{R}_+ \to \mathbb{R}$ is T_0-periodic or C^k we always mean in the time variable. Similar considerations hold for functions $f : E \times \mathcal{S}_{T_0} \to \mathbb{R}$. By means of this identification, we can rewrite (3.7) and (3.8) as

$$\mu^{(n)}(f) = \frac{1}{n}\int_0^{nT_0} f(X_t, t)\mathrm{d}t, \qquad f : V \times \mathcal{S}_{T_0} \to \mathbb{R}, \tag{3.9}$$

$$Q^{(n)}(g) = \frac{1}{n}\sum_{\substack{t \in [0, nT_0]: \\ X_{t-} \neq X_t}} g(X_{t-}, X_t, t), \qquad g : E \times \mathcal{S}_{T_0} \to \mathbb{R}, \tag{3.10}$$

where f, g are bounded and measurable.

To simplify the notation from now on we set

$$\mathcal{M}_* := \mathcal{M}_{+,T_0}(V \times \mathcal{S}_{T_0}) \times \mathcal{M}_+(E \times \mathcal{S}_{T_0}). \tag{3.11}$$

Definition 3.5. *We introduce the subset $\Lambda \subset \mathcal{M}_*$ given by the pairs $(\mu, Q) \in \mathcal{M}_*$ such that:*

(i) $\mu = \mu_t\, dt$ *with* $\mu_t(V) = 1$ *for almost every* $t \in \mathcal{S}_{T_0}$*;*
(ii) $Q = Q_t\, dt$*;*
(iii) $\partial_t \mu_t + \operatorname{div} Q_t = 0$ *weakly;*
(iv) *for almost every* $t \in \mathcal{S}_{T_0}$ *it holds:* $\mu_t(y) = 0 \Rightarrow Q_t(y,z) = 0$ *for all* $(y,z) \in E$.

Theorem 2. *Given $x \in V$ the family $\{\mathbb{P}_x \circ (\mu^{(n)}, Q^{(n)})^{-1}\}_{n \geq 1}$ of probability measures on $\mathcal{M}_*$ satisfies a large deviation principle with speed n and good and convex rate functional I defined as*

$$I(\mu, Q) := \begin{cases} \int_0^{T_0} I_t(\mu_t, Q_t)\, dt & \text{if } (\mu, Q) \in \Lambda, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.12}$$

The proof of the above theorem is given in Sects. 8 and 9.

We recall that the above LDP means that, for any $\mathcal{C} \subset \mathcal{M}_*$ closed and any $\mathcal{G} \subset \mathcal{M}_*$ open, it holds

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x\big((\mu^{(n)}, Q^{(n)}) \in \mathcal{C}\big) \leq - \inf_{(\mu,Q)\in\mathcal{C}} I(\mu, Q), \tag{3.13}$$

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x\big((\mu^{(n)}, Q^{(n)}) \in \mathcal{G}\big) \geq - \inf_{(\mu,Q)\in\mathcal{G}} I(\mu, Q). \tag{3.14}$$

Remark 3.6. We point out that $I(\mu, Q) = 0$ if and only if $\mu = \pi_t\, \mathrm{d}t$ and $Q = Q_t \mathrm{d}t$ with $Q_t(y,z) = \pi_t(y) r(y,z;t)$, $(y,z) \in E$. Indeed, by the properties of the function Φ stated after (3.1), $I_t(\mu_t, Q_t) = 0$ if and only if $Q_t(y,z) = \mu_t(y) r(y,z;t)$ for any $(y,z) \in E$. By Proposition 7.5 and the continuity equation, this last property is satisfied for almost all t only when $\mu = \pi_t\, \mathrm{d}t$ and $Q = Q_t \mathrm{d}t$ with $Q_t(y,z) = \pi_t(y) r(y,z;t)$ for any $(y,z) \in E$.

Since $\bar{\mu}_{nT_0}(\cdot) = \frac{1}{T_0}\mu^{(n)}(\cdot, \mathcal{S}_{T_0})$ and $\bar{Q}_{nT_0}(\cdot) = \frac{1}{T_0} Q^{(n)}(\cdot, \mathcal{S}_{T_0})$, Theorem 1 with T of the form nT_0 follows from Theorem 2 by applying the contraction principle.

3.3. LDP for Currents

Recalling that E denotes the set of ordered edges of V with strictly positive jump rates, we let $E_s := \{(y,z) \in V \times V \,:\, (y,z) \in E \text{ or } (z,y) \in E\}$ be the symmetrization of E in $V \times V$. We denote by $\mathbb{R}_a^{E_s}$ the family of functions $\bar{J} : E_s \to \mathbb{R}$ which are antisymmetric, i.e., $\bar{J}(y,z) = -\bar{J}(z,y)$ $\forall (y,z) \in E_s$.

Definition 3.7. *Given $T > 0$, to each path $X \in D(\mathbb{R}_+; V)$ we associate the* empirical current $\bar{J}_T(X) \in \mathbb{R}_a^{E_s}$ *defined as*

$$\bar{J}_T(X)(y,z) = \frac{1}{T} \sum_{\substack{t\in[0,T]:\\ X_{t-}\neq X_t}} \Big[\mathbb{1}\big((X_{t-},X_t)=(y,z)\big) - \mathbb{1}\big((X_{t-},X_t)=(z,y)\big) \Big] \tag{3.15}$$

for any $(y,z) \in E_s$.

To introduce the *extended empirical current*, we denote by $\mathcal{M}_a(E_s \times \mathcal{S}_{T_0})$ the space of signed measures J on $E_s \times \mathcal{S}_{T_0}$ which are antisymmetric in E_s (i.e., $J(y,z,A) = -J(z,y,A)$ for any $A \subset \mathcal{S}_{T_0}$ measurable) and have finite total variation (i.e., J can be written as difference of two measures in $\mathcal{M}_+(E_s \times \mathcal{S}_{T_0})$). $\mathcal{M}_a(E_s \times \mathcal{S}_{T_0})$ is endowed with the usual weak topology, i.e., $\nu_n \to \nu$ if and only if $\nu_n(f) \to \nu(f)$ for any continuous function on $E_s \times \mathcal{S}_{T_0}$.

Definition 3.8. *Given $T > 0$, to each path $X \in D(\mathbb{R}_+; V)$ we associate the* extended empirical current $J^{(n)}(X) \in \mathcal{M}_a(E_s \times \mathcal{S}_{T_0})$ *defined as*

$$\begin{aligned} &J^{(n)}(y,z,B) \\ &\quad = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\substack{t\in B+kT_0:\\ X_{t-}\neq X_t}} \Big[\mathbb{1}\big((X_{t-},X_t)=(y,z)\big) - \mathbb{1}\big((X_{t-},X_t)=(z,y)\big) \Big] \end{aligned} \tag{3.16}$$

for any $B \subset [0,T_0)$ measurable.

We introduce the continuous map

$$\mathcal{J} : \mathcal{M}_+(E \times \mathcal{S}_{T_0}) \to \mathcal{M}_a(E_s \times \mathcal{S}_{T_0}) \tag{3.17}$$

defined as

$$\mathcal{J}(Q)(y,z,A) := Q(y,z,A) - Q(z,y,A),$$

for any $(y,z) \in E_s$ and $A \subset \mathcal{S}_{T_0}$ measurable, with the convention that $Q(y',z',A) := 0$ if $(y',z') \notin E$. Trivially, the following relation holds between the extended empirical flow and current:

$$\mathcal{J}\big(Q^{(n)}\big) = J^{(n)}. \tag{3.18}$$

As a consequence, from the contraction principle and the joint LDP for $(\mu^{(n)}, Q^{(n)})$ given in Theorem 2, we get that a joint LDP holds for $(\mu^{(n)}, J^{(n)})$ with speed n and good and convex rate functional $\widehat{I}(\mu, J)$ given by

$$\widehat{I}(\mu,J) := \inf_{Q:\mathcal{J}(Q)=J} I(\mu,Q). \tag{3.19}$$

It turns out that the above variational problem expressing the new rate functional $\widehat{I}$ can be exactly solved, thus leading to Theorem 3 below. In order to state this theorem, we need a preliminary definition:

Definition 3.9. *The set Λ_a is given by the pairs $(\mu, J) \in \mathcal{M}_{+,T_0}(V \times \mathcal{S}_{T_0}) \times \mathcal{M}_a(E_s \times \mathcal{S}_{T_0})$ such that*

(i) $\mu = \mu_t\,dt$ *with* $\mu_t(V) = 1$ *for almost every* $t \in \mathcal{S}_{T_0}$*;*
(ii) $J = J_t\,dt$*;*
(iii) $\partial_t \mu_t + \mathrm{div}\,J_t = 0$ *where* $\mathrm{div}\,J_t(y) = \sum_{(y,z)\in E_s} J_t(y,z)$*;*
(iv) *for almost every* $t \in \mathcal{S}_{T_0}$ *it holds:* $\mu_t(y) = 0 \Rightarrow J_t(y,z) \le 0$ *for all* $(y,z) \in E_s$*;*
(v) $J_t(y,z) \ge 0$ *if* $(y,z) \in E$ *and* $(z,y) \notin E$*, while* $J_t(y,z) \le 0$ *if* $(y,z) \notin E$ *and* $(z,y) \in E$

We recall that the continuity equation in Item (iii) has to be thought in its weak form. To state the joint LDP for $(\mu^{(n)}, J^{(n)})$ we introduce also the function $\Psi : \mathbb{R}\times\mathbb{R}\times\mathbb{R}_+ \mapsto [0,+\infty]$ given by[1]

$$\Psi(u,\bar{u};a) := \begin{cases} u\left[\operatorname{arcsinh}\frac{u}{a} - \operatorname{arcsinh}\frac{\bar{u}}{a}\right] \\ \qquad - \left[\sqrt{a^2+u^2} - \sqrt{a^2+\bar{u}^2}\right] & \text{if } a>0, \\ \Phi(|u|,|\bar{u}|) & \text{if } a=0. \end{cases} \tag{3.20}$$

We recall that $\operatorname{arcsinh}(x) = \log[x+\sqrt{x^2+1}]$.

Finally, for the theorem below, we recall that $r(y,z;t) := 0$ if $(y,z) \notin E$.

Theorem 3. *Given* $x \in V$ *the family* $\left\{\mathbb{P}_x \circ (\mu^{(n)}, J^{(n)})^{-1}\right\}_{n\ge 1}$ *of probability measures on*

$$\mathcal{M}_{+,T_0}(V\times\mathcal{S}_{T_0}) \times \mathcal{M}_a(E_s\times\mathcal{S}_{T_0})$$

satisfies a large deviation principle with speed n *and good and convex rate functional* $\widehat{I}$ *given by*

$$\widehat{I}(\mu,J) = \begin{cases} \int_0^{T_0} I_t(\mu_t, Q_t^{J,\mu})\,dt & \text{if } (\mu,J)\in\Lambda_a, \\ +\infty & \text{otherwise,} \end{cases} \tag{3.21}$$

where

$$Q_t^{J,\mu}(y,z) = \frac{J_t(y,z) + \sqrt{J_t^2(y,z) + 4\mu_t(y)\mu_t(z)r(y,z;t)r(z,y;t)}}{2}. \tag{3.22}$$

Moreover, given $(\mu,J)\in\Lambda_a$*, the rate functional* $\widehat{I}(\mu,J)$ *can be rewritten as*

$$\widehat{I}(\mu,J) = \frac{1}{2}\sum_{(y,z)\in E_s}\int_0^{T_0} \Psi\left(J_t(y,z), J_t^\mu(y,z); a_t^\mu(y,z)\right)dt, \tag{3.23}$$

where

$$J_t^\mu(y,z) := \mu_t(y)r(y,z;t) - \mu_t(z)r(z,y;t),$$
$$a_t^\mu(y,z) := 2\sqrt{\mu_t(y)\mu_t(z)r(y,z;t)r(z,y;t)}.$$

The proof of Theorem 3 is given in Sect. 10.

[1] This formula corresponds to [6, Eq. (6.3)], apart the correction of a typo there in the case $a = 0$.

Remark 3.10. By contraction, Theorem 3 implies a joint LDP for the empirical measure and current. The corresponding convex rate functional $\bar{\widehat{I}} : \mathcal{P}(V) \times \mathbb{R}_a^{E_s} \to [0, +\infty]$ is given by

$$\bar{\widehat{I}}(\bar{\mu}, \bar{J}) = \inf_{(\mu,J)} \frac{1}{T_0} \widehat{I}(\mu, J), \tag{3.24}$$

where the infimum is taken among all pairs (μ, J) in Λ_a such that $\frac{1}{T_0} \int_0^{T_0} \mu_t \mathrm{d}t = \bar{\mu}$ and $\frac{1}{T_0} \int_0^{T_0} J_t \mathrm{d}t = \bar{J}$, where $\mu = \mu_t \mathrm{d}t$ and $J = J_t \mathrm{d}t$.

4. Stochastic Entropy Flow

In this section, we assume that $(y, z) \in E$ if and only if $(z, y) \in E$, i.e., $E = E_s$.

One usually defines the *fluctuating entropy flow* on the time interval $[0, nT_0]$ as the Radon–Nikodym derivative

$$\sigma_{nT_0}[X] = \log \left. \frac{d\mathbb{P}}{d\left(\mathbb{P}^B \circ R_{nT_0}\right)} \right|_{[0,nT_0]} \left((X_s)_{s\in[0,nT_0]} \right), \tag{4.1}$$

where $\mathbb{P}|_{[0,nT_0]}$ is the law on $D([0, nT_0]; V)$ of the continuous-time Markov chain with rates $r(\cdot, \cdot; t)$ and some initial distribution μ_0 and $\mathbb{P}^B\big|_{[0,nT_0]}$ is the law on $D([0, nT_0]; V)$ of another continuous-time Markov chain with rates $r^B(\cdot, \cdot; t)$, and some initial distribution μ_0'. Then, the measure $\mathbb{P}^B \circ R_{nT_0}$ is the pushforward measure of the law $\mathbb{P}^B$ in the time window $[0, nT_0]$ by R_{nT_0}, R_{nT_0} being the pathways time reversal in the time window $[0, nT_0]$. Of course, definition (4.1) restricts to the case when the Radon–Nikodym derivative is well defined. We further restrict to the case of T_0-periodic rates $r(\cdot, \cdot; t)$ and $r^B(\cdot, \cdot; t)$. Below we will consider only three peculiar choices of rates $r^B(\cdot, \cdot; t)$: the naive reversal (cf. Sect. 4.1), the reversed protocol first used in [15] (cf. Sect. 4.2) and the dual reversed protocol first considered for time inhomogeneous processes in [15] (cf. Sect. 4.3). The fact that by playing with different choices of the backward process (i.e., the rates r^B here) we find different physical quantities (excess heat, housekeeping heat, total heat, phase-space contraction,...) has been pointed out in [10,14] for diffusion processes and in [32] for pure jump Markov processes. We recall that the excess heat and the housekeeping heat were first introduced by Oono and Paniconi [53]: the excess heat measures the non-stationarity of the process, while the housekeeping heat measures the distance of the process from the "instantaneous" reversibility. These two quantities permit to obtain a refinement of the second law of thermodynamics (see [10,14,32]). The total heat is then the sum of the excess heat and the housekeeping heat. On the other hand, the phase-space contraction is a quantity characterizing the irreversibility of a deterministic dynamical system [27,63]. For diffusion processes, see [14] for the fluctuation relation associated with the phase-space contraction and to its generalization to obtain the fluctuation relation of the finite time Lyapunov exponents.

We point out that (4.1) implies directly the finite time fluctuation relation

$$\mathbb{P}\left(\sigma_{nT_0}[X] \in [\sigma, \sigma + d\sigma)\right) \exp(-\sigma) = \mathbb{P}^B\left(\sigma^B_{nT_0}[X] \in [-\sigma, -\sigma + d\sigma)\right), \quad (4.2)$$

with the backward entropy flow $\sigma^B_{nT_0}[X]$ defined by

$$\sigma^B_{nT_0}[X] = \log \left.\frac{d\mathbb{P}^B}{d\left(\mathbb{P} \circ R_{nT_0}\right)}\right|_{[0,nT_0]} \big((X_s)_{s\in[0,nT_0]}\big). \quad (4.3)$$

By using Eq. (7.1) in Sect. 7.1 and the periodicity of the rates, we get

$$\begin{aligned}\sigma_{nT_0}[X] = \log \frac{\mu_0(X_0)}{\mu_0'(X_{nT_0})} &+ \sum_{\substack{s\in(0,nT_0]:\\ X_{s-}\neq X_s}} \log \frac{r(X_{s-}, X_s; s)}{r^B(X_s, X_{s-}; T_0 - s)} \\ &- \int_0^{nT_0} \mathrm{d}s \left[r(X_s; s) - r^B(X_s; T_0 - s)\right].\end{aligned} \quad (4.4)$$

We point out that since V is finite, the boundary term b.t. $= \log \frac{\mu_0(X_0)}{\mu_0'(X_{nT_0})}$ will not play any role in the large deviation limit. We now provide three examples of entropy flows relevant in statistical physics and show that—apart negligible boundary terms—they can be expressed by contraction from the empirical extended measure and/or flow.

4.1. Entropy Flow From Naive Reversal

We take the *identity reversal* $r^B(y, z; t) := r(y, z; t)$, and we write $\sigma^{\text{naive}}_{nT_0}$ for the associated entropy flow. We define the functional $S_{\text{naive}}(\mu, Q)$ on the space $\mathcal{M}_*$ introduced in (3.11) as follows:

$$\begin{aligned}S_{\text{naive}}(\mu, Q) := &\sum_{(y,z)\in E} \int Q(y, z, \mathrm{d}s) \log \frac{r(y, z; s)}{r(z, y; T_0 - s)} \\ &- \sum_y \int \mu(y, \mathrm{d}s)[r(y; s) - r(y, T_0 - s)].\end{aligned} \quad (4.5)$$

Then, the entropy flow fulfills the identity

$$\frac{1}{n}\sigma^{\text{naive}}_{nT_0} = \frac{\text{b.t.}}{n} + S_{\text{naive}}\big(\mu^{(n)}, Q^{(n)}\big). \quad (4.6)$$

4.2. Total Entropy Flow from Reversed Protocol

As rates r^B (denoted here by r^{R}), we choose the *reversed protocol*, i.e., we take the rates

$$r^{\mathrm{R}}(y, z; t) := r(y, z; T_0 - t). \quad (4.7)$$

The resulting entropy flow, usually called *total entropy flow*, will be denoted by $\sigma^{\text{tot}}_{nT_0}$. Defining the functional S_{tot} as

$$S_{\text{tot}}(Q) := \sum_{(y,z)\in E} \int Q(y, z, \mathrm{d}s) \log \frac{r(y, z; s)}{r(z, y; s)} \quad (4.8)$$

for $Q \in \mathcal{M}_+(E \times \mathcal{S}_{T_0})$, we get

$$\frac{1}{n}\sigma^{\rm tot}_{nT_0} = \frac{\rm b.t.}{n} + S_{\rm tot}\big(Q^{(n)}\big). \tag{4.9}$$

The above entropy production has been investigated in [57,59] for time periodic processes.

4.3. Entropy Flow in Excess

Given t we write w_t for the *accompanying distribution* (cf. [33]), which is defined as the unique invariant distribution of the time homogeneous (and continuous time) Markov chain with frozen jump rates $r(\cdot,\cdot;t)$. Due to our assumptions (A1) and (A2), the distribution w_t is well defined, and moreover it is strictly positive on each state of V.

As rates r^B (denoted here by $r^{\rm DR}$), we then choose the *dual reversed protocol*:

$$r^{\rm DR}(y,z;t) = w^{-1}_{T_0-t}(y)r(z,y;T_0-t)w_{T_0-t}(z). \tag{4.10}$$

The resulting entropy flow, denoted by $\sigma^{\rm ex}_{nT_0}$ and called *excess entropy flow*, is related to the excess heat discussed in [34,53]. By the invariance of w_t, from (4.4) one easily gets the simplified expression

$$\sigma^{\rm ex}_{nT_0}[X] = {\rm b.t.} + \sum_{\substack{s\in(0,nT_0]:\\ X_{s-}\neq X_s}} \log\frac{w_s(X_s)}{w_s\big(X_{s-}\big)}. \tag{4.11}$$

At the cost of a boundary term (irrelevant in the LD limit) we find

$$\sigma^{\rm ex}_{nT_0}[X] = {\rm b.t.}' - \int_0^{nT_0} {\rm d}s\,[\partial_s(\log w_s)](X_s)\,, \tag{4.12}$$

which is the quantity considered in the time periodic setup by Schuller et al. in [64]. By defining the functional

$$S_{\rm ex}(\mu) := -\sum_y \int \mu(y,{\rm d}s)\partial_s \log\big(w_s(y)\big) \tag{4.13}$$

for $\mu \in \mathcal{M}_{+,T_0}(V\times\mathcal{S}_{T_0})$, we can write

$$\frac{1}{n}\sigma^{\rm ex}_{nT_0} = \frac{b.t.'}{n} + S_{\rm ex}\big(\mu^{(n)}\big). \tag{4.14}$$

5. Gallavotti–Cohen Duality Relations

As in Sect. 4 we assume that $(y,z)\in E$ if and only if $(z,y)\in E$, i.e., $E=E_s$.

We recall that, given $(\mu,Q)\in\mathcal{M}_*$, it holds $I(\mu,Q)=+\infty$ if $(\mu,Q)\notin\Lambda$ (see (3.11) and (3.12)). Hence, for the analysis of Gallavotti–Cohen duality relations, we restrict to $(\mu,Q)\in\Lambda$.

Definition 5.1. *Given $(\mu,Q)\in\Lambda$, with $\mu=\mu_t dt$ and $Q=Q_t dt$, we define the transformed element $(\theta\mu,\theta Q)\in\Lambda$ as $\theta\mu=(\theta\mu_t)dt$, $\theta Q=(\theta Q_t)dt$ where*

$$\theta\mu_t := \mu_{T_0-t}, \qquad \theta Q_t(y,z) := Q_{T_0-t}(z,y). \tag{5.1}$$

It is simple to check that $(\theta\mu, \theta Q)$ is indeed an element of Λ.

In what follows, we write $I(\mu, Q; r)$ for the joint LD rate functional of Theorem 2 referred to the Markov chain with jump rates $r(y, z; t)$. Similarly, we add the reference to the jump rates in the entropy production functions by writing $S_{\text{naive}}(\mu, Q; r)$, $S_{\text{tot}}(Q; r)$ and $S_{\text{ex}}(\mu; r)$ (recall the notation introduced in Sects. 4.1, 4.2 and 4.3). By means of the contraction principle, one derives from Theorem 2 the LDP for the entropy production functions $S_{\text{naive}}(\mu, Q; r)$, $S_{\text{tot}}(Q; r)$ and $S_{\text{ex}}(\mu; r)$ with LD functionals given, respectively, by

$$\begin{aligned}\mathcal{I}_{\text{naive}}(s; r) &= \inf\{I(\mu, Q; r) \,:\, S_{\text{naive}}(\mu, Q; r) = s\},\\ \mathcal{I}_{\text{tot}}(s; r) &= \inf\{I(\mu, Q; r) \,:\, S_{\text{tot}}(Q; r) = s\},\\ \mathcal{I}_{\text{ex}}(s; r) &= \inf\{I(\mu, Q; r) \,:\, S_{\text{ex}}(\mu; r) = s\}.\end{aligned}$$

Theorem 4. *For any $(\mu, Q) \in \Lambda$, we have the following level 2.5 Gallavotti–Cohen duality relations:*

$$I(\theta\mu, \theta Q; r) = I(\mu, Q; r) + S_{\text{naive}}(\mu, Q; r), \tag{5.2}$$

$$I(\theta\mu, \theta Q; r^{\text{R}}) = I(\mu, Q; r) + S_{\text{tot}}(Q; r), \tag{5.3}$$

$$I(\theta\mu, \theta Q; r^{\text{DR}}) = I(\mu, Q; r) + S_{\text{ex}}(\mu; r). \tag{5.4}$$

Moreover, for any real s we have by contraction the following Gallavotti–Cohen duality relations:

$$\mathcal{I}_{\text{naive}}(-s; r) = \mathcal{I}_{\text{naive}}(s; r) + s, \tag{5.5}$$

$$\mathcal{I}_{\text{tot}}(-s; r^{\text{R}}) = \mathcal{I}_{\text{tot}}(s; r) + s, \tag{5.6}$$

$$\mathcal{I}_{\text{ex}}(-s; r^{\text{DR}}) = \mathcal{I}_{\text{ex}}(s; r) + s. \tag{5.7}$$

The above duality relations are new with exception of (5.6) which appears also in [59,69]. The proof of Theorem 4 is given in Sect. 11.

If we have a time symmetric protocol, i.e., $r\,(y, z; T_0 - t) = r\,(y, z; t)$, then the naive entropy flow and the total entropy flow are identical and the duality relations (5.5) and (5.6) become identical. If the accompanying distribution satisfies the instantaneous detailed balance such that the relation (4.10) becomes $r^{DR}\,(y, z; t) = r\,(y, z; T_0 - t)$, then the excess entropy flow and the total entropy flow are identical. In particular, the duality relations (5.6) and (5.7) become identical. Finally, we point out that in [64] the Gallavotti–Cohen relation has been experimentally checked in a context where the two previous situations both take place; hence, in that context the three duality relations (5.5), (5.6) and (5.7) are identical.

By the contraction principle, the duality relations in Theorem 4 imply some analogous relations for the extended current. To this aim, we define $\theta J_t(y, z) = J_{T_0-t}(z, y) = -J_{T_0-t}(y, z)$ and write $\widehat{I}(\mu, J; r)$ for the LD rate functional $\widehat{I}(\mu, J)$ of $(\mu^{(n)}, J^{(n)})$ with jump rates $r(\cdot, \cdot; \cdot)$ (see (3.19) and (3.23)). In particular, one derives the following corollary:

Corollary 5.2. *For any* $(\mu, J) \in \Lambda_a$ *it holds*

$$\widehat{I}(\theta\mu, \theta J; r^{\mathrm{R}}) - \widehat{I}(\mu, J; r) = \frac{1}{2} \sum_{(y,z)\in E} \int_0^{T_0} J_t(y,z) \log \frac{r(y,z;t)}{r(z,y;t)}\, dt, \tag{5.8}$$

$$\widehat{I}(\theta\mu, \theta J; r^{\mathrm{DR}}) - \widehat{I}(\mu, J; r) = S_{\mathrm{ex}}(\mu; r). \tag{5.9}$$

Proof. Recall the map $\mathcal{J}$ defined in (3.17) and note that

$$\mathcal{J}(Q) = J \Longrightarrow \mathcal{J}(\theta Q) = \theta J. \tag{5.10}$$

Observe also that, if $(\mu, Q) \in \Lambda$ is such that $\mathcal{J}(Q) = J$, then

$$S_{\mathrm{tot}}(Q; r) = \frac{1}{2} \sum_{(y,z)\in E} \int_0^{T_0} J_t(y,z) \log \frac{r(y,z;t)}{r(z,y;t)}\, \mathrm{d}t. \tag{5.11}$$

The duality relation (5.8) then follows from (3.19) by taking the infimum in both sides of (5.3) among all Q with $(\mu, Q) \in \Lambda$ and $\mathcal{J}(Q) = J$, and by using (5.10) and (5.11). The duality relation (5.9) follows by the same procedure applied to (5.4). □

We remark that as the naive entropy flow (4.5) cannot be expressed (up to boundary terms) as contraction of the extended empirical measure and current, there is no version of Corollary 5.2 (i.e., with extended current) for the duality relation (5.2). Finally, by applying once again the contraction principle to Corollary 5.2 we get other duality relations (we omit the proof since simple):

Corollary 5.3. *It holds*

$$I_c(\theta J; r^{\mathrm{R}}) - I_c(J; r) = \frac{1}{2} \sum_{(y,z)\in E} \int_0^{T_0} J_t(y,z) \log \frac{r(y,z;t)}{r(z,y;t)}\, dt, \tag{5.12}$$

$$I_m(\theta\mu; r^{\mathrm{DR}}) - I_m(\mu; r) = S_{\mathrm{ex}}(\mu; r), \tag{5.13}$$

where I_c *and* I_m *denote the LD rate functionals of the extended empirical current and the extended empirical measure, respectively.*

The first relation is the Gallavotti–Cohen relation for the LD rate functional of the extended empirical current only, and the second relation is a level 2-duality relation for the LD rate functional of the extended empirical density only. We are not aware of previously derived relation of the type of (5.13) even in time homogeneous setup.

Finally, we point out that the LD rate functional $\bar{I}(\bar{\mu}, \bar{Q})$ (cf. (3.3)) and $\bar{\bar{I}}(\bar{\mu}, \bar{J})$ (cf. (3.24)) do not satisfy duality relations resulting from a naive contraction of the relations in Theorem 4. Indeed, the three entropy flows cannot be expressed as contraction of the empirical measure and empirical flow/current (recall Definitions 3.1 and 3.7).

6. Two-State Systems

We consider the simplest possible system, that is, a two-state ($V = \{0,1\}$) chain. In this case, the model is completely determined by the two periodic functions $r_t(0,1)$ and $r_t(1,0)$ that fix the jump rates. (For simplicity of notation, sometimes the time variable t will appear as subindex in the rates.) Even if elementary, this framework has, however, interesting and non-trivial physical applications.

We list some relevant examples:

- In [57] we have a quantum dot with one single active energy level periodically modulated that corresponds to a two-state Markov chain with rates
$$\begin{cases} r_t(0,1) = \frac{\Gamma}{1+\exp(x_t)}, \\ r_t(1,0) = \frac{\Gamma \exp(x_t)}{1+\exp(x_t)}, \end{cases} \tag{6.1}$$
where x_t is time periodic and related to the energy of the quantum dot, the chemical potential and the temperature of the bath.
- In [64] we have a single defect center in natural IIa-type diamond excited by a red and a green laser with time periodic intensity. The corresponding rates are
$$\begin{cases} r_t(0,1) = a_0\big(1 + \gamma \sin\big(\frac{2\pi}{T_0} t\big)\big), \\ r_t(1,0) = b_0. \end{cases} \tag{6.2}$$
- In [50], we have a two-state model of stochastic resonance given by
$$\begin{cases} r_t(0,1) = \exp\big(-k \cos\big(\frac{2\pi}{T_0} t\big)\big), \\ r_t(1,0) = \exp\big(k \cos\big(\frac{2\pi}{T_0} t\big)\big). \end{cases} \tag{6.3}$$
- In [70], it is discussed a piecewise constant and symmetric protocol
$$\begin{cases} r_t(0,1) = \exp\big(-h_t\big), \\ r_t(1,0) = \exp\big(h_t\big), \end{cases} \quad \text{with } h_t = \begin{cases} h_0 - a & \text{if } 0 \le t \le \alpha T_0, \\ h_0 + a & \text{if } \alpha T_0 \le t \le T_0, \end{cases} \tag{6.4}$$
for some $\alpha \in (0,1)$.

Let us now discuss some results concerning the general situation. In all this section we restrict to elements μ, Q, J with $(\mu, Q) \in \Lambda$ and $(\mu, J) \in \Lambda_a$, without further mention. For convenience, we call $\mu_t := \mu_t(0)$, $Q_t := Q_t(0,1)$ and $J_t := J_t(0,1)$ (note that this is different from the usual notation); accordingly, the jump rates are here denoted by $r_t(0,1)$ and $r_t(1,0)$. The continuity equation is simply $\partial_t \mu_t + J_t = 0$. Note that, by the above continuity equation, the knowledge of μ_t and Q_t allows to recover

$$\mu_t(1) = 1 - \mu_t(0), \qquad Q_t(1,0) = \partial_t \mu_t + Q_t, \qquad J_t = -\partial_t \mu_t. \tag{6.5}$$

On the other hand, given real functions μ_t and Q_t defined for $t \in \mathcal{S}_{T_0}$ and setting (6.5), we have that $(\mu, Q) \in \Lambda$ if and only if $\mu_t \in [0,1]$, $Q_t \ge 0$ and $\partial_t \mu_t + Q_t \ge 0$.

Moreover (recall that we restrict to $(\mu, Q) \in \Lambda$), the LD rate functional of Theorem 2 becomes

$$I(\mu, Q) = \int_0^{T_0} \Big[Q_t \log \frac{Q_t}{\mu_t r_t(0,1)} + (\partial_t \mu_t + Q_t) \log \frac{(\partial_t \mu_t + Q_t)}{(1-\mu_t) r_t(1,0)} \\ + \mu_t r_t(0,1) + (1-\mu_t) r_t(1,0) - 2Q_t \Big] \mathrm{d}t. \tag{6.6}$$

In this case, one can compute explicitly the LD rate functional $I_m(\mu) = \inf_Q I(\mu, Q)$ associated with the extended empirical measure $\mu^{(n)}$. We have that $I_m(\mu)$ coincides in this case with the joint LD functional for measure and current, i.e., $I_m(\mu) = \widehat{I}(\mu, J)$. This is because the current is completely determined by the density using $\partial_t \mu_t = -J_t$. (This fact is indeed true for more general Markov chains, and indeed it is enough that the undirected graph obtained from the transition graph by disregarding the orientation and identifying multiple edges is a tree.) The rate functional $I_m(\mu)$ is therefore obtained as

$$I_m(\mu) = I(\mu, Q(\mu, \partial_t \mu)), \tag{6.7}$$

where (cf. (3.22))

$$Q_t(\mu, \partial_t \mu) := \frac{-\partial_t \mu_t + \sqrt{(\partial_t \mu_t)^2 + 4\mu_t(1-\mu_t) r_t(0,1) r_t(1,0)}}{2}. \tag{6.8}$$

We point out that, in general, given $(\mu, Q) \in \Lambda$ the level 2.5 rate functional $I(\mu, Q)$ is the time integration of $I_t(\mu_t, Q_t)$ and the latter is related to the level 2.5 rate functional (for the non-extended empirical measure and flow) with frozen jump rates $r(\cdot, \cdot; t)$. One could wonder if the same property holds for the level 2 rate functional $I_m(\mu)$. In particular, for 2-state Markov chains, one could wonder if $I_m(\mu)$ equals

$$\int_0^{T_0} I_t^{\text{frozen}}(\mu_t) \,\mathrm{d}t = \int_0^{T_0} \left(\sqrt{\mu_t r_t(0,1)} - \sqrt{(1-\mu_t) r_t(1,0)} \right)^2 \mathrm{d}t. \tag{6.9}$$

Formula (6.9) follows by the explicit form of the level 2 rate functional for a 2-state chain, which is always reversible [19–22]. This property (i.e., the identity between $I_m(\mu)$ and (6.9)) does not hold in general. Indeed, since by (6.8)

$$Q_t(\mu, 0) = \sqrt{\mu_t(1-\mu_t) r_t(0,1) r_t(1,0)} \tag{6.10}$$

it holds

$$I_t^{\text{frozen}}(\mu_t) = I_t(\mu_t, Q_t(\mu, 0)), \tag{6.11}$$

where $I_t(\cdot, \cdot)$ denotes the integrand in the r.h.s. of (6.6). The above identities (6.11) and (6.7) imply that $I_m(\mu)$ equals (6.9) when μ is constant in time (even with time-dependent rates), and imply that the zeroth-order term of the formal expansion in $\partial_t \mu$ of $I_m(\mu) = I(\mu, Q(\mu, \partial_t \mu))$ coincides with (6.9).

In [70] the LD rate functional of the excess entropy flow (called there "cumulated work") for a two-state model with a time symmetric piecewise constant protocol is computed explicitly (cf. Equation (20) there). This explicit level 1 LD rate functional could be obtained by the contraction from our previous formulas.

The 2-state case is simple enough to allow also an explicit computation of the non-equilibrium oscillatory state π. By a direct computation, we have

$$\pi_t(0) = \frac{e^{-\Gamma_t}}{1-e^{-\Gamma_{T_0}}}\left[\int_0^t r_s(1,0)e^{\Gamma_s}\,\mathrm{d}s + e^{-\Gamma_{T_0}}\int_t^{T_0} r_s(1,0)e^{\Gamma_s}\,\mathrm{d}s\right],$$
$$\pi_t(1) = \frac{e^{-\Gamma_t}}{1-e^{-\Gamma_{T_0}}}\left[\int_0^t r_s(0,1)e^{\Gamma_s}\,\mathrm{d}s + e^{-\Gamma_{T_0}}\int_t^{T_0} r_s(0,1)e^{\Gamma_s}\,\mathrm{d}s\right],$$

where $\Gamma_t := \int_0^t [r_s(0,1)+r_s(1,0)]\,\mathrm{d}s$. Indeed, it is simple to verify that $\pi_t(0) \geq 0$, $\pi_t(1) \geq 0$, $\pi_t(0)+\pi_t(1) = 1$ and that the continuity equation, which reduces to $\partial_t \pi_t(0) + \pi_t(0)r_t(0,1) - \pi_t(1)r_1(1,0) = 0$, is fulfilled. Note that [25, Prop. 3.13] provides an alternative formula for π_t. Recall that $I(\mu,Q)$ is zero when $\mu_t(y) = \pi_t(y)$ and $Q_t(y,z) = \pi_t(y)r_t(y,z)$.

From now on we restrict to the special case $r_t := r_t(0,1) = r_t(1,0)$. In this case, it is possible to obtain an explicit expression for the rate functional $\bar{I}_f\left(\bar{Q}\right)$ of the empirical flow $\bar{Q}_T$ when $T \to +\infty$ (see Remark 3.2). By the graphical construction, since the jump rates are the same, we have that $\bar{Q}_T$ coincides up to negligible terms with $\frac{\mathcal{N}_T}{2T}$ where $\mathcal{N}_T$ is a non-homogeneous Poisson process with periodic intensity given by r_t. When $T = nT_0$, we can write $\mathcal{N}_T = \sum_{i=1}^n Y_i$, where the Y_i are i.i.d Poisson random variables of parameter $\int_0^{T_0} r_t \mathrm{d}t$. The variable Y_i represents the number of points in the interval $((i-1)T_0, iT_0]$. Using the classic Cramér's theorem, we deduce that

$$\bar{I}_f(\bar{Q}) = 2\bar{Q}\log\left[\frac{2\bar{Q}}{\bar{r}}\right] - 2\bar{Q} + \bar{r}, \qquad \bar{r} := \frac{1}{T_0}\int_0^{T_0} r_t \mathrm{d}t. \tag{6.12}$$

The above result can be also obtained variationally by showing that the minimizer in

$$\bar{I}_f\left(\bar{Q}\right) := \frac{1}{T_0}\inf_{\left\{(\mu,Q):\,\frac{1}{T_0}\int_0^{T_0} Q_t \mathrm{d}t = \bar{Q}\right\}} I(\mu,Q), \tag{6.13}$$

is given by $\mu_t = \frac{1}{2}$ and $Q_t = r_t\bar{Q}/\bar{r}$. We omit the computations.

Comparison with an Effective Time Homogenous Chain

Always in the case of equal jump rates, i.e., $r_t := r_t(0,1) = r_t(1,0)$, we here obtain an upper bound for the rate functional $\bar{I}(\bar{\mu},\bar{Q})$ (see (3.3)) in terms of the level 2.5 rate functional of a time homogenous Markov chain with suitable rates (in the same spirit of homogenization theory).

Let us call $I^{\bar{r}}$ the LD rate functional for the empirical measure and flow of a 2-state Markov chain having time independent rates equal to $r(0,1) = r(1,0) = \bar{r}$. According to [5,6], we have

$$I^{\bar{r}}(\bar{\mu},\bar{Q}) = \bar{Q}\log\left[\frac{\bar{Q}^2}{\bar{\mu}(0)\bar{\mu}(1)\bar{r}^2}\right] - 2\bar{Q} + \bar{r}, \tag{6.14}$$

where, by the divergence-free condition, $\bar{Q} := \bar{Q}(0,1) = \bar{Q}(1,0)$. (Also below, we restrict to divergence-free flows $\bar{Q}$; otherwise, we have $I^{\bar{r}}(\bar{\mu}, \bar{Q}) = \infty$.) By minimizing (6.14) among $\bar{\mu}$ and comparing with (6.12), we get that

$$\inf_{\bar{\mu}} I^{\bar{r}}(\bar{\mu}, \bar{Q}) = \bar{I}_f(\bar{Q}) = \inf_{\bar{\mu}} \bar{I}(\bar{\mu}, \bar{Q}).$$

In addition, we can show the inequality

$$\bar{I}(\bar{\mu}, \bar{Q}) \leq I^{\bar{r}}(\bar{\mu}, \bar{Q}), \tag{6.15}$$

which in general is strict. Inequality (6.15) can be derived simply by inserting in (3.3) the special pair (μ, Q) given by

$$\mu_t(y) = \bar{\mu}(y), \qquad Q_t(y,z) = \frac{r_t(y,z)\bar{Q}(y,z)}{\bar{r}}.$$

Considering more general Markov chains, one cannot expect inequality (6.15) to be true. Indeed, such an inequality would imply that the rate functionals have the same global minima, which in general is not valid; see Remark 3.3.

7. Preliminary Results

In this section, we collect some technical results. Since some of them will be applied also to a tilted continuous-time Markov chain with less regular jump rates; here, we only assume that the jump rates satisfy the periodicity assumption (i.e., $r(\cdot,\cdot;t) = r(\cdot,\cdot;t+T_0)$ for some $T_0 > 0$), assumptions (A1) and (A2) and that $r(y,z;\cdot)$ is a measurable, locally integrable nonnegative function. As mentioned in Sect. 2, the last assumption guarantees that the associated continuous-time Markov chain is well defined [17].

Definition 7.1. *Given $\mu \in \mathcal{M}_{+,T_0}(V \times \mathcal{S}_{T_0})$ we define $Q^\mu \in \mathcal{M}_+(E \times \mathcal{S}_{T_0})$ as $Q^\mu(y,z,dt) := \mu(y,dt)r(y,z;t)$. If $\mu = \mu_t dt$, then we set $Q^\mu_t(y,z) := \mu_t(y)r(y,z;t)$ (thus implying that $Q^\mu = Q^\mu_t\, dt$).*

7.1. Radon–Nikodym Derivative

Calling N_t the number of jumps of the trajectory X up to time t, and $\tau_1 < \tau_2 < \cdots < \tau_{N_t}$ the jump times, then it holds for $0 < t_1 < t_2 < \cdots < t_n < t$ and $x_1, x_2, \ldots, x_n \in V$

$$\begin{aligned} &\mathbb{P}_x \begin{pmatrix} N_t = n, \tau_i \in [t_i, t_i + \mathrm{d}t_i) \text{ for all } i = 1,2,\ldots,n, \\ \xi_t = x_i \text{ for all } t \in [t_i, t_{i+1}), \forall i = 0,1,\ldots,n \end{pmatrix} \\ &= \exp\Big\{ - \sum_{i=0}^{n} \int_{t_i}^{t_{i+1}} r(x_i; s)\mathrm{d}s \Big\} \prod_{i=0}^{n-1} r(x_i, x_{i+1}; t_{i+1}) \mathrm{d}t_1 \mathrm{d}t_2 \ldots \mathrm{d}t_n, \end{aligned} \tag{7.1}$$

where $x_0 := x$, $t_0 := 0$ and $t_{n+1} := t$. We recall that $r(x;s) := \sum_z r(x,z;s)$.

We consider another Markov chain on V with T_0-periodic rates $\bar{r}(y,z;t)$ (given by nonnegative locally integrable functions) and such that

$$\bar{r}(y,z;t) > 0 \qquad \Longrightarrow \qquad r(y,z;t) > 0.$$

Then, its law $\bar{\mathbb{P}}_x|_{[0,t]}$ on the space $D([0,t];V)$ of càdlàg paths is absolutely continuous with respect to $\mathbb{P}_x|_{[0,t]}$ and the Radon–Nikodym derivative on $D([0,t];V)$ is given by

$$\frac{d\bar{\mathbb{P}}_x}{d\mathbb{P}_x}\Big|_{[0,t]}\big((X_s)_{s\in[0,t]}\big) = \exp\left\{\int_0^t \big[r(X_s;s)-\bar{r}(X_s;s)\big]\mathrm{d}s\right\}\prod_{\substack{s\in(0,t]:\\ X_{s-}\neq X_s}}\frac{\bar{r}(X_{s-},X_s;s)}{r(X_{s-},X_s;s)}. \tag{7.2}$$

Let us suppose that $r(y,z;t)=0$ if and only if $\bar{r}(y,z;t)=0$. Then, we can write

$$\bar{r}(y,z;t)=r(y,z;t)e^{F(y,z;t)}, \qquad F(y,z;t):=\log\frac{\bar{r}(y,z;t)}{r(y,z;t)}$$

(above we used the convention $\log(0/0)=0$). Note that F is T_0–periodic. Since $r(y;\cdot)$ and $r(y,z;\cdot)$ are T_0-periodic functions, we can restate (7.2) as follows:

$$\frac{d\mathbb{P}_x^F}{d\mathbb{P}_x}\Big|_{[0,nT_0]} = \exp\Big\{n\mu^{(n)}(r-\bar{r})+nQ^{(n)}(F)\Big\}, \qquad \mathbb{P}_x^F:=\bar{\mathbb{P}}_x. \tag{7.3}$$

7.2. Some Identities

Take $Q\in\mathcal{M}_+(E\times\mathcal{S}_{T_0})$. Denoting by $\mathcal{B}$ the Borel sets of $\mathcal{S}_{T_0}$, for each $y\in V$

$$\mathcal{B}\ni A\mapsto\sum_z Q(y,z,A)-\sum_z Q(z,y,A)=:\operatorname{div}Q(y,A)\in\mathbb{R}$$

is a signed measure on $\mathcal{S}_{T_0}$. In what follows, we denote by $\operatorname{div}Q(f)$ the integral of f w.r.t. the above measure $\operatorname{div}Q$:

$$\operatorname{div}Q(f)=\sum_y\int_0^{T_0}\operatorname{div}Q(y,\mathrm{d}s)f(y,s), \qquad f:V\times\mathcal{S}_{T_0}\to\mathbb{R}. \tag{7.4}$$

Lemma 7.2. *Let $f:V\times\mathcal{S}_{T_0}\to\mathbb{R}$ be C^1. Then,*

$$\mu^{(n)}(\partial_s f)-\operatorname{div}Q^{(n)}(f)=\frac{1}{n}\big(f(X_{nT_0},0)-f(X_0,0)\big). \tag{7.5}$$

Proof. Let $s_1<s_2<\cdots<s_m$ be the jump times of the path X in the time interval $(kT_0,(k+1)T_0]$. We set $s_0:=kT_0$ and $s_{m+1}:=(k+1)T_0$. We can write

$$\begin{aligned} f(X_{(k+1)T_0},(k+1)T_0)-f(X_{kT_0},kT_0) &= f(X_{s_m},(k+1)T_0)-f(X_{s_1-},kT_0)\\ &=\sum_{j=0}^m[f(X_{s_j},s_{j+1})-f(X_{s_j},s_j)]+\sum_{j=1}^m[f(X_{s_j},s_j)-f(X_{s_j-},s_j)]\\ &=\sum_{j=0}^m\int_{s_j}^{s_{j+1}}\partial_s f(X_s,s)\mathrm{d}s+\sum_{j=1}^m[f(X_{s_j},s_j)-f(X_{s_j-},s_j)]\\ &=\int_{kT_0}^{(k+1)T_0}\partial_s f(X_s,s)\mathrm{d}s+\sum_{j=1}^m[f(X_{s_j},s_j)-f(X_{s_j-},s_j)]. \end{aligned}$$

Averaging the above identities among $k = 0, \dots, n-1$ and using the T_0-periodicity of f we get

$$\frac{1}{n}\big(f(X_{nT_0},0) - f(X_0,0)\big) = \mu^{(n)}(\partial_s f) + \sum_{y,z}\int_{[0,T_0]} Q^{(n)}(y,z,\mathrm{d}s)\big(f(z,s) - f(y,s)\big)$$
$$= \mu^{(n)}(\partial_s f) - \sum_y\sum_z\int_{[0,T_0]} Q^{(n)}(y,z,\mathrm{d}s)f(y,s) + \sum_y\sum_z\int_{[0,T_0]} Q^{(n)}(z,y,\mathrm{d}s)f(y,s)$$
$$= \mu^{(n)}(\partial_s f) - \operatorname{div} Q^{(n)}(f).$$

□

7.3. The Oscillatory Steady State

We collect in the following proposition some asymptotic properties of the oscillatory steady state. Recall the definition of π given in Sect. 2 and Definition 7.1.

Proposition 7.3. *The following holds:*

(i) *Fixed $t \in [0,T_0]$, under $\mathbb{P}_x$, the law of X_{t+nT_0} weakly converges to π_t as n goes to ∞;*

(ii) *$\mathbb{P}_x$-a.s. $\mu^{(n)}$ weakly converges to $\pi = \pi_t dt$ in $\mathcal{M}_{+,T_0}(V \times \mathcal{S}_{T_0})$. More generally, given a measurable function $f : V \times \mathcal{S}_{T_0} \to \mathbb{R}$ with $\|f\|_\infty < \infty$, it holds*

$$\lim_{n\to\infty} \mu^{(n)}(f) = \pi(f) \qquad \mathbb{P}_x\text{-a.s. and in } L^1(\mathbb{P}_x). \tag{7.6}$$

(iii) *$\mathbb{P}_x$-a.s., $\bar{\mu}_T$ weakly converges to $\frac{1}{T_0}\int_0^{T_0} \pi_t dt$ in $\mathcal{P}(V)$;*

(iv) *$\mathbb{P}_x$-a.s. $Q^{(n)}(y,z;dt)$ weakly converges to $Q^\pi_t(y,z)dt$ in $\mathcal{M}_+(E \times \mathcal{S}_{T_0})$. More generally, given a measurable function $g : E\times\mathcal{S}_{T_0} \to \mathbb{R}$ with $\|g\|_\infty < \infty$, it holds*

$$\lim_{n\to\infty} Q^{(n)}(g) = Q^\pi(g) \qquad \mathbb{P}_x\text{-a.s. and in } L^1(\mathbb{P}_x). \tag{7.7}$$

Proof. (i) Due to Assumptions (A1) and (A2), the discrete-time Markov chain $(X_{t+nT_0})_{n\geq 0}$ is irreducible. Since V is finite, we get that this discrete-time Markov chain has a unique invariant distribution to which it converges (whatever the initial distribution). As a consequence, the invariant distribution must be given by the distribution π_t introduced in Sect. 2. This concludes the proof of Item (i). Item (iii) follows directly from Item (ii). The proof of Items (ii) and (iv) can be derived from [35, Thm. 2.1] adapted to processes with càdlàg paths. We comment this step. We associate with the continuous-time Markov chain $\xi = (\xi_t)_{t\in\mathbb{R}_+}$ the random sequence $\mathbb{X} = (\mathbb{X}_k)_{k\geq 0}$ of paths in $D([0,T_0];V)$ with $\mathbb{X}_k := (\xi_{kT_0+s})_{0\leq s\leq T_0}$. By the arguments presented to derive Theorem 2.1 in [35], we get that $\mathbb{X}$ is a Markov chain, ergodic and stationary when ξ_0 is sampled with distribution π_0. Hence, $\mathbb{P}_{\pi_0}|_{[0,T_0]}$ is the marginal distribution of $\mathbb{X}$ in the stationary state. Given measurable functions $f : V \times \mathcal{S}_{T_0} \to \mathbb{R}$ and $g : E \times \mathcal{S}_{T_0} \to \mathbb{R}$ with $\|f\|_\infty, \|g\|_\infty < \infty$, we can write (see (3.9) and (3.10))

$$\mu^{(n)}(f) = \frac{1}{n}\sum_{j=0}^{n-1} F(\mathbb{X}_j), \qquad F(\zeta) := \int_0^{T_0} f(\zeta_t,t)\mathrm{d}t, \tag{7.8}$$

$$Q^{(n)}(g) = \frac{1}{n}\sum_{j=0}^{n-1} G(\mathbb{X}_j) + O(1/n), \qquad G(\zeta) := \sum_{\substack{t\in[0,T_0):\\ \zeta_{t-}\neq\zeta_t}} g(\zeta_{t-},\zeta_t,t), \tag{7.9}$$

where ζ denotes a generic element of $D([0,T_0];V)$. We observe that F,G are integrable w.r.t. $\mathbb{P}_{\pi_0}|_{[0,T_0]}$. This is trivial for F since bounded. The integrability of G follows from the boundedness of g and the fact that the total number of jumps in $[0,T_0)$ under $\mathbb{P}_{\pi_0}|_{[0,T_0]}$ is stochastically dominated by a suitable Poisson random variable due to Assumption (A3); hence $G \in L^1(\mathbb{P}_{\pi_0})$. From Birkhoff's ergodic theorem, we derive the $\mu^{(n)}(f)$ converges to $\mathbb{E}_{\pi_0}[F]$ and $Q^{(n)}(g)$ converges to $\mathbb{E}_{\pi_0}[G]$ both $\mathbb{P}_{\pi_0}$–a.s. and in $L^1(\mathbb{P}_{\pi_0})$. Since $\mathbb{P}_{\pi_0} = \sum \pi_0(x)\mathbb{P}_x$ and $\pi_0(x) > 0$ for any x, we derive the convergence also $\mathbb{P}_x$–a.s. and in $L^1(\mathbb{P}_x)$ for any $x \in V$. □

Lemma 7.4. *It holds $\partial_t \pi_t + \operatorname{div} Q^\pi_t = 0$ weakly.*

Proof. Due to Definition 7.1, we only need to prove that $\pi(\partial_s f) - \operatorname{div} Q^\pi(f) = 0$ for any C^1 function $f : V \times \mathcal{S}_{T_0} \to \mathbb{R}$. This identity can be obtained by taking the limit $n \to \infty$ in Lemma 7.2 and using Proposition 7.3. □

We conclude this section with an alternative characterization of $\pi = \pi_t \mathrm{d}t$.

Proposition 7.5. *The only weak solution $\mu \in \mathcal{M}_{+,T_0}(V \times \mathcal{S}_{T_0})$, with $\mu = \mu_t\, dt$, of the equation*

$$\partial_t \mu_t + \operatorname{div} Q^\mu_t = 0 \tag{7.10}$$

is given by π.

Proof. We first show that $\mu \in \mathcal{M}_{+,T_0}(V \times \mathcal{S}_{T_0})$, with $\mu = \mu_t \mathrm{d}t$, solving (7.10) is an invariant measure of the piecewise-deterministic Markov process $(W_t, Y_t)_{t\geq 0}$ on $V \times \mathcal{S}_{T_0}$ with extended generator [17] given by (2.1) (we are making some slight abuse of notation, since (x,s) in (2.1) has to be thought of as element of $V \times \mathcal{S}_{T_0}$ via the canonical projection for times). By [17, Thm. (26.14)], the domain of the extended generator is given by the functions $f(x,s)$ which are absolutely continuous in s (we shortly write $f \in \mathcal{AC}$). Hence, due to [17, Prop. (34.7)], μ is an invariant measure for the PDMP if and only in $\mu(Lf) = 0$ for any $f \in \mathcal{AC}$. By density, it is enough that $\mu(Lf) = 0$ for any C^1 function f, which (by integration by parts) is equivalent to the fact that μ is a weak solution of (7.10).

We set $p_{s,s+t}(x,y) := P(\xi_{t+s} = y|\xi_s = x)$ for the probability transition of the Markov chain. Since μ is an invariant measure for the PDMP, given a C^1 function f on $V \times \mathcal{S}_{T_0}$, it holds

$$\begin{aligned}\sum_x \int_{\mathcal{S}_{T_0}} \mu_s(x) f(x,s)\mathrm{d}s &= \sum_x \int_{\mathcal{S}_{T_0}} \mu_s(x) E\big[f(W_{T_0}, Y_{T_0})|W_0 = x, Y_0 = s\big]\mathrm{d}s \\ &= \sum_x \sum_y \int_{\mathcal{S}_{T_0}} \mu_s(x) p_{s,s+T_0}(x,y) f(y, s+T_0)\mathrm{d}s\end{aligned}$$

$$= \sum_y \sum_x \int_{\mathcal{S}_{T_0}} \mu_s(y) p_{s,s+T_0}(y,x) f(x,s) \mathrm{d}s. \tag{7.11}$$

Note that, in the last identity, we have used the T_0–periodicity of f. By the T_0–periodicity of the map $s \mapsto \mu_s$ and by the arbitrariness of f in (7.11), we conclude that $\mu_{s+T_0}(x) = \mu_s(x) = \sum_y \mu_s(y) p_{s,s+T_0}(y,x)$. This is the equation characterizing π, apart a multiplicative factor. As a consequence, we get that $\mu = c\pi$ for some factor c. Since both μ and π have total mass T_0, we conclude that $\mu = \pi$. On the other hand, it is simple to check (by the arguments presented above) that $\mu := \pi$ solves (7.10). □

8. Proof of Theorem 2: upper bound (3.13), convexity and lower-semicontinuity of I

We start by showing exponential tightness:

Lemma 8.1. *The family* $\{\mathbb{P}_x \circ (\mu^{(n)}, Q^{(n)})^{-1}\}_{n \geq 1}$ *of probability measures on* $\mathcal{M}_*$ *is exponentially tight.*

Proof. Given $\ell > 0$ we set $\mathcal{K}_\ell := \{(\mu, Q) \in \mathcal{M}_* : Q(1) \leq \ell\}$. Above, $Q(1)$ denotes the averaged value w.r.t. to the measure Q of the function constantly equal to 1, equivalently $Q(1)$ is the total mass of the measure Q. Then, $\mathcal{K}_\ell$ is a compact subset of $\mathcal{M}_*$ [2]. To prove the exponential tightness, it is enough to show that there exists $C > 0$ such that

$$\varlimsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}_x(\, (\mu^{(n)}, Q^{(n)}) \notin \mathcal{K}_\ell\,) \leq -C\ell \tag{8.1}$$

for large ℓ.

We prove (8.1). The event $\{(\mu^{(n)}, Q^{(n)}) \notin \mathcal{K}_\ell\}$ is simply the event that the measure $Q^{(n)}$ has total mass larger than ℓ. Due to (3.10), the total mass of $Q^{(n)}$ equals $1/n$ times the number of jumps in the time interval $[0, nT_0]$. On the other hand, by the graphical construction presented in Sect. 2.1 the number of jumps in the time interval $[0, nT_0]$ is stochastically dominated by a Poisson variable Z of parameter $\lambda n T_0$ where $\lambda = \sum_{(y,z)} \sup_{t \in [0,T_0]} r(y,z;t)$. Since $E[e^{\gamma Z}] = \exp\{\lambda n T_0 (e^\gamma - 1)\}$, by applying Chebyshev's inequality, we get

$$\begin{aligned} \mathbb{P}_x(\, (\mu^{(n)}, Q^{(n)}) \notin \mathcal{K}_\ell\,) &= \mathbb{P}_x(\, Q^{(n)}(1) > \ell\,) \leq P(Z > n\ell) \\ &\leq e^{-n\ell} E[e^Z] = \exp\{-n\ell + \lambda n T_0 (e-1)\}. \end{aligned} \tag{8.2}$$

The above bound trivially implies (8.1). □

Recall that $r(y;t) = \sum_z r(y,z;t)$. Given a continuous function $F : E \times \mathcal{S}_{T_0} \to \mathbb{R}$, we set

$$r^F(y,z;t) = r(y,z;t) e^{F(y,z;t)} \qquad \text{and} \qquad r^F(y;t) = \sum_z r^F(y,z;t).$$

Moreover, we consider $\phi : V \times \mathcal{S}_{T_0} \to \mathbb{R}$ of class C^1 and we define the mappings $\widehat{I}_{\phi,F} : \mathcal{M}_* \to \mathbb{R}_+$ and $I_{\phi,F} : \mathcal{M}_* \to [0,+\infty]$ as follows:

$$\widehat{I}_{\phi,F}(\mu,Q) := -\mu(\partial_t \phi) + \operatorname{div} Q(\phi) + Q(F) - \mu(r^F - r), \tag{8.3}$$

$$I_{\phi,F}(\mu,Q) := \begin{cases} \widehat{I}_{\phi,F}(\mu,Q) & \text{if } \mu = \mu_t \mathrm{d}t,\ \mu_t(V) = 1 \text{ a.s.} \\ +\infty & \text{otherwise.} \end{cases} \tag{8.4}$$

Lemma 8.2. *The function $I_{\phi,F}$ is convex and lower semicontinuous.*

Proof. Let us call $\mathcal{A}$ the set of pairs $(\mu,Q) \in \mathcal{M}_*$ such that $\mu = \mu_t \mathrm{d}t$, $\mu_t(V) = 1$ for almost all $t \in \mathcal{S}_{T_0}$. It is simple to check that $\mathcal{A}$ is convex and closed in $\mathcal{M}_*$. Since $\mathcal{A}$ is convex and $\widehat{I}_{\phi,F}$ is convex, it is simple to derive that $I_{\phi,F}(\mu,Q)$ is convex.

Let us now prove that $I_{\phi,F}$ is continuous on $\mathcal{A}$. To this aim, given $(\nu^{(k)}, Q^{(k)}) \to (\nu,Q)$ in $\mathcal{A}$, we need to show that $\widehat{I}_{\phi,F}(\nu^{(k)},Q^{(k)}) \to \widehat{I}_{\phi,F}(\nu,Q)$. Due to the definition of weak convergence of measures and since $\partial_t\phi$, ϕ and F are continuous, the only non-trivial step is to show that $\nu^{(k)}(h) \to \nu(h)$ where $h := r^F - r$. Since $h(y;t) = \sum_z r(y,z;t)[e^{F(y,z;t)} - 1]$, and F is continuous in time, for each y the function $h(y;\cdot)$ is continuous on $\mathcal{S}_{T_0} \setminus \mathcal{D}$ (recall Assumption (A4)). On the other hand, since $\nu = \nu_t \mathrm{d}t$, we have $\sum_y \nu(y,\mathcal{D}) = 0$. As a by-product of the last observation and the Portmanteau theorem as stated in [51, Thm. 12.6], we get that $\nu^{(k)}(h) \to \nu(h)$.

This concludes the proof that $I_{\phi,F}$ is continuous on the set $\mathcal{A}$. Since $I_{\phi,F}$ is continuous on the closed set $\mathcal{A}$ and it equals $+\infty$ on $\mathcal{M}_* \setminus \mathcal{A}$, we conclude that $I_{\phi,F}$ is lower semicontinuous. □

Let us define

$$M_n^F := \exp\Big\{ - n\mu^{(n)}(r^F - r) + nQ^{(n)}(F) \Big\}. \tag{8.5}$$

We recall that by (7.3)

$$\frac{d\mathbb{P}_x^F}{d\mathbb{P}_x}\Big|_{[0,nT_0]} = M_n^F$$

where $\mathbb{P}_x^F$ is the law of the new Markov chain with jump rates $r^F(y,z;t)$.

Due to (7.5), we can write

$$- nI_{\phi,F}(\mu^{(n)},Q^{(n)}) = \phi(X_{nT_0},0) - \phi(X_0,0) - \log M_n^F\,. \tag{8.6}$$

In the above identity, we have used also that $\mu^{(n)}(x,\mathrm{d}t) = \mu_t^{(n)}(x)\mathrm{d}t$ where $0 \leq \mu_t^{(n)}(x) \leq 1$ (cf. (3.7)), thus implying that $I_{\phi,F}(\mu^{(n)},Q^{(n)}) = \widehat{I}_{\phi,F}(\mu^{(n)},Q^{(n)})$.

Lemma 8.3. *Fix $x \in V$. For each ϕ, F as above and each measurable $\mathcal{B} \subset \mathcal{M}_*$ it holds*

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x\Big((\mu^{(n)},Q^{(n)}) \in \mathcal{B}\Big) \leq - \inf_{(\mu,Q)\in\mathcal{B}} I_{\phi,F}(\mu,Q). \tag{8.7}$$

Proof. Due to (8.6), we can write

$$\begin{aligned}
&\mathbb{P}_x\left((\mu^{(n)}, Q^{(n)}) \in \mathcal{B}\right) \\
&\quad = \mathbb{E}_x\big(\exp\big\{-nI_{\phi,F}(\mu^{(n)}, Q^{(n)}) - [\phi(X_{nT_0},0) - \phi(X_0,0)]\big\} M_n^F \mathbb{1}_{\mathcal{B}}(\mu^{(n)}, Q^{(n)})\big) \\
&\quad \leq \big[\sup_{(\mu,Q)\in\mathcal{B}} e^{-nI_{\phi,F}(\mu,Q)}\big] e^{2\|\phi\|_\infty} \mathbb{E}_x(M_n^F) = \big[\sup_{(\mu,Q)\in\mathcal{B}} e^{-nI_{\phi,F}(\mu,Q)}\big] e^{2\|\phi\|_\infty},
\end{aligned}$$

thus implying the thesis. □

Due to the exponential tightness (see Lemma 8.1), it is enough to prove the upper bound (3.13) for compact subsets $\mathcal{K} \subset \mathcal{M}_*$ instead of generic closed subsets $\mathcal{C} \subset \mathcal{M}_*$. Due to Lemma 8.3, for any open subset $\mathcal{O} \subset \mathcal{M}_*$ we have

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x\left((\mu^{(n)}, Q^{(n)}) \in \mathcal{O}\right) \leq -\sup_{\phi,F} \inf_{(\mu,Q)\in\mathcal{O}} I_{\phi,F}(\mu,Q).$$

As a by-product of the above bound and the minmax lemma (cf. [41, Lemma 3.3, App. 2]), we conclude that

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x\left((\mu^{(n)}, Q^{(n)}) \in \mathcal{K}\right) \leq -\inf_{(\mu,Q)\in\mathcal{K}} \sup_{\phi,F} I_{\phi,F}(\mu,Q).$$

Hence, to conclude the proof of the upper bound (3.13) it is enough to apply the following lemma:

Lemma 8.4. *For each $(\mu, Q) \in \mathcal{M}_*$, it holds*

$$I(\mu,Q) = \sup_{\phi,F} I_{\phi,F}(\mu,Q), \tag{8.8}$$

where the supremum is taken among all C^1 functions $\phi : V \times \mathcal{S}_{T_0} \to \mathbb{R}$ and continuous functions $F : E \times \mathcal{S}_{T_0} \to \mathbb{R}$.

Remark 8.5. Note that, by the convexity and the lower-semicontinuity of $I_{\phi,F}$, the above lemma implies the convexity and the lower-semicontinuity of I.

Proof. In what follows, we write $\ell(\cdot)$ for the Lebesgue measure on $\mathcal{S}_{T_0}$.

- *Case* $(\mu, Q) \notin \Lambda$. We claim that (8.8) reduces to $+\infty = +\infty$ if $(\mu, Q) \notin \Lambda$.

From the definition of $I(\mu, Q)$ and $I_{\phi,F}(\mu, Q)$ (see (3.12) and (8.4)), one trivially gets that both sides of (8.8) are $+\infty$ if $(\mu, Q) \notin \mathcal{A}$, where $\mathcal{A}$ is defined as in the proof of Lemma 8.2. It is also trivial to verify that both sides of (8.8) are $+\infty$ if, for some C^1 function $\phi : V \times \mathcal{S}_{T_0} \to \mathbb{R}$, it holds $-\mu(\partial_t\phi) + \operatorname{div} Q(\phi) \neq 0$. Hence, in what follows we restrict to the case $\mu = \mu_t \mathrm{d}t$, $\mu_t(V) = 1$ for almost all $t \in \mathcal{S}_{T_0}$, and $\partial_t\mu + \operatorname{div} Q = 0$ (in the weak sense). Since in this case $I_{\phi,F}(\mu, Q)$ does not depend on ϕ, we write simply $I_F(\mu, Q)$.

Suppose now that Q is not of the form $Q_t\mathrm{d}t$. Hence, there exists a subset B of $\mathcal{S}_{T_0}$ with zero Lebesgue measure such that $Q(y_0, z_0, B) > 0$ for some $(y_0, z_0) \in E$. Since both $\ell(\cdot)$ and $Q(y_0, z_0, \cdot)$ are measures of finite mass, they are regular. Hence, by [2, Thm. 1.1], for any $\varepsilon > 0$ there exist a closed set D_ε and an open set G_ε such that $D_\varepsilon \subset B \subset G_\varepsilon$, $Q(y_0, z_0, G_\varepsilon \setminus D_\varepsilon) \leq \varepsilon$ and $\ell(G_\varepsilon \setminus D_\varepsilon) \leq \varepsilon$. In what follows, we take $\varepsilon < Q(y_0, z_0, B)/2$, thus implying that $Q(y_0, z_0, D_\varepsilon) \geq Q(y_0, z_0, B)/2$. On the other hand, since $\ell(B) = 0$, we get that $\ell(G_\varepsilon) \leq \varepsilon$. By Urysohn's lemma we can find a continuous function $\varphi_\varepsilon : \mathcal{S}_{T_0} \to [0,1]$ such that $\varphi_\varepsilon \equiv 1$ on D_ε and $\varphi_\varepsilon \equiv 0$ on G_ε^c. We then introduce the continuous test function $F_\varepsilon(y,z,t) = \gamma(\varepsilon)\delta_{y,y_0}\delta_{z,z_0}\varphi_\varepsilon(t)$ where the positive parameter $\gamma(\varepsilon)$ will be fixed at the end. Then, we have

$$\begin{aligned}
&I_{F_\varepsilon}(\mu, Q)\\
&= \sum_{(y,z)} \int_0^{T_0} Q(y,z,\mathrm{d}t)F_\varepsilon(y,z,t) - \sum_y \int_0^{T_0} \mu_t(y)\big(r^{F_\varepsilon}(y,t) - r(y,t)\big)\mathrm{d}t\\
&= \int_0^{T_0} Q(y_0,z_0,\mathrm{d}t)F_\varepsilon(y_0,z_0,t) - \int_0^{T_0} \mu_t(y_0)r(y_0,z_0,t)\big(e^{F_\varepsilon(y_0,z_0,t)} - 1\big)\mathrm{d}t\\
&\geq \gamma(\varepsilon)Q(y_0,z_0,D_\varepsilon) - e^{\gamma(\varepsilon)} \int_0^{T_0} \mu_t(y_0)r(y_0,z_0,t)\mathbb{1}(t \in G_\varepsilon)\mathrm{d}t\\
&\geq \gamma(\varepsilon)Q(y_0,z_0,D_\varepsilon) - e^{\gamma(\varepsilon)}\ell(G_\varepsilon) \max_{y,z,t} r(y,z,t)\\
&\geq \gamma(\varepsilon)Q(y_0,z_0,B)/2 - e^{\gamma(\varepsilon)}\varepsilon \max_{y,z,t} r(y,z,t). \qquad (8.9)
\end{aligned}$$

Taking $\gamma(\varepsilon) := \log(1/\varepsilon)$, we get that $\lim_{\varepsilon\downarrow 0} I_{F_\varepsilon}(\mu, Q) = +\infty$. Hence, it holds $\sup_F I_F(\mu, Q) = +\infty$, while trivially $I(\mu, Q) = +\infty$ since $(\mu, Q) \notin \Lambda$.

We now focus on property (iv) in Definition 3.5 of Λ. Let us suppose that there exist $B \subset \mathcal{S}_{T_0}$ and an edge (y_0, z_0) such that $\ell(B) > 0$, $\mu_t(y_0) = 0$ for all $t \in B$ and $Q_t(y_0, z_0) > 0$ for all $t \in B$. We need to prove that $\sup_F I_F(\mu, Q) = \infty$. As above for any $\varepsilon > 0$ we fix a closed set D_ε and an open set G_ε such that $D_\varepsilon \subset B \subset G_\varepsilon$ and $\ell(G_\varepsilon \setminus D_\varepsilon) \leq \varepsilon$. Without loss of generality, we take $D_\varepsilon \subset D_{\varepsilon'}$ if $\varepsilon > \varepsilon'$. Since $\ell(B) > 0$, we have $\ell(D_{\varepsilon_0}) \geq \ell(B)/2 > 0$ for $\varepsilon_0 := \ell(B)/2$. In particular, $\int_{D_\varepsilon} Q_t(y_0, z_0)\mathrm{d}t \geq \int_{D_{\varepsilon_0}} Q_t(y_0, z_0)\mathrm{d}t > 0$ for any $\varepsilon < \varepsilon_0$. Hence, similarly to (8.9), we get

$$I_{F_\varepsilon}(\mu, Q) \geq \gamma(\varepsilon) \int_{D_{\varepsilon_0}} Q_t(y_0, z_0)\mathrm{d}t - e^{\gamma(\varepsilon)}\ell(G_\varepsilon \setminus B) \max_{y,z,t} r(y,z,t).$$

Using that $\ell(G_\varepsilon \setminus D_\varepsilon) \leq \varepsilon$ and taking $\gamma(\varepsilon) := \log(1/\varepsilon)$, we conclude that $\lim_{\varepsilon\downarrow 0} I_{F_\varepsilon}(\mu, Q) = +\infty$, thus proving that $\sup_F I_F(\mu, Q) = \infty$.

This concludes the proof of our initial claim.

• *Case* $(\mu, Q) \in \Lambda$. We now assume that $(\mu, Q) \in \Lambda$. Since $\partial_t \phi + \operatorname{div} Q = 0$, we have $I_{\phi,F}(\mu, Q) = I_{0,F}(\mu, Q) =: I_F(\mu, Q)$. Hence, we only need to show that

$$I(\mu, Q) = \sup_F I_F(\mu, Q), \qquad (8.10)$$

where the supremum is taken among the continuous functions $F : E\times\mathcal{S}_{T_0} \to \mathbb{R}$ and

$$\begin{aligned} I_F(\mu,Q) &= \sum_{(y,z)} \int_0^{T_0} Q_t(y,z)F(y,z,t)\mathrm{d}t - \sum_y \int_0^{T_0} \mu_t(y)\big(r^F(y;t) - r(y;t)\big)\mathrm{d}t \\ &= \sum_{(y,z)} \int_0^{T_0} \mathrm{d}t\Big[Q_t(y,z)F(y,z,t) - \mu_t(y)r(y,z;t)(e^{F(y,z,t)} - 1)\Big]. \end{aligned} \tag{8.11}$$

Since (cf. (3.1)) $\Phi(q,p) = \sup_{v\in\mathbb{R}}\{qv - p(e^v - 1)\}$ for any $(q,p) \in \mathbb{R}_+ \times \mathbb{R}_+$, we can bound from above the integrand in the r.h.s. of (8.11) by $\Phi\big(Q_t(y,z), \mu_t(y)r(y,z;t)\big)$, thus implying that

$$I_F(\mu,Q) \le \sum_{(y,z)} \int_0^{T_0} \Phi\big(Q_t(y,z), \mu_t(y)r(y,z;t)\big)\mathrm{d}t = I(\mu,Q). \tag{8.12}$$

It remains to prove that $I(\mu,Q) \le \sup_F I_F(\mu,Q)$, F varying among the continuous functions. Since $(\mu,Q) \in \Lambda$ we have

$$I(\mu,Q) = \sum_{(y,z)\in E} \int_0^{T_0} \Phi\big(Q_t(y,z), \mu_t(y)r(y,z;t)\big)\mathrm{d}t.$$

Given $(y,z) \in E$ and given $\varepsilon > 0$ we define

$$\begin{aligned} A(y,z) &:= \{t \in \mathcal{S}_{T_0} \,:\, Q_t(y,z) = 0\}, \\ B(y,z) &:= \{t \in \mathcal{S}_{T_0} \,:\, \mu_t(y) = 0 \text{ and } Q_t(y,z) > 0\}, \\ C(y,z) &:= \mathcal{S}_{T_0} \setminus \big(A(y,z) \cup B(y,z)\big) = \{t \in \mathcal{S}_{T_0} \,:\, Q_t(y,z) > 0 \text{ and } \mu_t(y) > 0\}, \\ C_\varepsilon(y,z) &:= \left\{t \in \mathcal{S}_{T_0} \,:\, \varepsilon \le Q_t(y,z) \le \frac{1}{\varepsilon} \text{ and } \varepsilon \le \mu_t(y) \le \frac{1}{\varepsilon}\right\}. \end{aligned}$$

Since $(\mu,Q) \in \Lambda$, we have $\ell\big(B(y,z)\big) = 0$. In particular, by definition of Φ (cf. (3.1)),

$$\begin{aligned} &\int_0^{T_0} \Phi\big(Q_t(y,z), \mu_t(y)r(y,z;t)\big)\mathrm{d}t \\ &\quad = \int_{A(y,z)} \mu_t(y)r(y,z;t)\mathrm{d}t + \int_{C(y,z)} \Phi\big(Q_t(y,z), \mu_t(y)r(y,z;t)\big)\mathrm{d}t. \end{aligned} \tag{8.13}$$

Since $\Phi\big(Q_t(y,z), \mu_t(y)r(y,z;t)\big) \in \mathbb{R}_+$ on $C(y,z)$ we have

$$\begin{aligned} &\int_{C(y,z)} \Phi\big(Q_t(y,z), \mu_t(y)r(y,z;t)\big)\mathrm{d}t \\ &\qquad = \lim_{\varepsilon\downarrow 0} \int_{C_\varepsilon(y,z)} \Phi\big(Q_t(y,z), \mu_t(y)r(y,z;t)\big)\mathrm{d}t. \end{aligned} \tag{8.14}$$

We now note that

$$\sum_{(y,z)\in E} \int_{C_\varepsilon(y,z)} \Phi\big(Q_t(y,z), \mu_t(y)r(y,z;t)\big)\mathrm{d}t = I_{F_\varepsilon}(\mu,Q), \tag{8.15}$$

where

$$F_\varepsilon(y,z,t) := \begin{cases} \log \frac{Q_t(y,z)}{\mu_t(y)r(y,z;t)} & \text{if } t \in C_\varepsilon(y,z), \\ 0 & \text{otherwise.} \end{cases}$$

Given $M > 0$ we now define

$$F_{M,\varepsilon}(y,z,t) := \begin{cases} -M & \text{if } \ t \in A(y,z), \\ \log \frac{Q_t(y,z)}{\mu_t(y)r(y,z;t)} & \text{if } \ t \in C_\varepsilon(y,z), \\ 0 & \text{otherwise.} \end{cases}$$

Note that $F_{M,\varepsilon}$ is a bounded measurable function. Since $\int_{A(y,z)} \mu_t(y)r(y,z;t)\mathrm{d}t$ $= \lim_{M\uparrow\infty} \int_{A(y,z)}(1-e^{-M})\mu_t(y)r(y,z;t)\mathrm{d}t$, from (8.13), (8.14) and (8.15) we conclude that

$$I(\mu,Q) = \sum_{(y,z)\in E} \int_0^{T_0} \Phi\big(Q_t(y,z),\mu_t(y)r(y,z;t)\big)\mathrm{d}t = \lim_{\varepsilon\downarrow 0} I_{F_{1/\varepsilon,\varepsilon}}(\mu,Q).$$

Above, we have used the same notation as in (8.11), which remains meaningful for bounded measurable functions. To have (8.10) it is now enough to approximate $I_{F_{1/\varepsilon,\varepsilon}}(\mu,Q)$ by $I_F(\mu,Q)$ with F continuous, for any fixed $\varepsilon > 0$. To this aim, we recall that by construction $F_{1/\varepsilon,\varepsilon}$ is a bounded measurable function. Let ψ_n be a sequence of continuous mollifiers. Then, $G_{n,\varepsilon}$ defined as the convolution of $F_{1/\varepsilon,\varepsilon}$ with ψ_n is a continuous function with $\|G_{n,\varepsilon}\|_\infty \leq \|F_{1/\varepsilon,\varepsilon}\|_\infty$ and such that $G_{n,\varepsilon} \to F_{1/\varepsilon,\varepsilon}$ Lebesgue almost everywhere. By (8.11) and dominated convergence, we then conclude that $\lim_{n\to\infty} I_{G_{n,\varepsilon}}(\mu,Q) = I_{F_{1/\varepsilon,\varepsilon}}(\mu,Q)$. ☐

9. Proof of Theorem 2: Lower Bound (3.14) and Goodness of *I*

The goodness of the rate functional follows from the exponential tightness in Lemma 8.1 and Lemma 4.1.23 in [18]. Our strategy to prove the lower bound is based on a relative entropy calculation according to the following general result, where $\mathrm{Ent}(\cdot|\cdot)$ denotes the relative entropy of probability distributions.

Lemma 9.1. *Let $\{P_n\}$ be a sequence of probability measures on a Polish space $\mathcal{X}$. Assume that for each $x \in \mathcal{X}$ there exists a sequence of probability measures $\{\widetilde{P}_n^x\}$ weakly convergent to δ_x and such that*

$$\varlimsup_{n\to\infty} \frac{1}{n} \mathrm{Ent}\left(\widetilde{P}_n^x \middle| P_n\right) \leq J(x) \tag{9.1}$$

for some $J\colon \mathcal{X} \to [0,+\infty]$. Then, the sequence $\{P_n\}$ satisfies the large deviation lower bound with rate functional given by $\mathrm{sc}^- J$, the lower semicontinuous envelope of J, i.e.,

$$\left(\mathrm{sc}^- J\right)(x) := \sup_{U\in\mathcal{N}_x} \inf_{y\in U} J(y)$$

where $\mathcal{N}_x$ denotes the collection of the open neighborhoods of x.

This lemma has been originally proven in [36, Prop. 4.1]; see also [48, Prop. 1.2.4].

We first prove the inequality (9.1) for the functional J defined as follows. Let $\Lambda_0 \subseteq \Lambda$ be the collection of elements $(\mu, Q) \in \Lambda$ such that there exists $\varepsilon > 0$ for which $\mu_t(x) \geq \varepsilon$ and $\varepsilon^{-1} \geq Q_t(y,z) \geq \varepsilon$ for all $t \in \mathcal{S}_{T_0}$, $x \in V$ and $(y,z) \in E$. We define

$$J(\mu, Q) = \begin{cases} I(\mu, Q) & \text{if } (\mu, Q) \in \Lambda_0, \\ +\infty & \text{otherwise.} \end{cases}$$

Then, we finish the proof of the lower bound showing that $(\mathrm{sc}^- J) = I$.

Given $(\mu, Q) \in \Lambda_0$, we consider a Markov chain $\widetilde{\mathbb{P}}$ having jump rates defined by

$$\widetilde{r}(y,z;t) := \frac{Q_t(y,z)}{\mu_t(y)}. \tag{9.2}$$

We observe that $\varepsilon \leq \widetilde{r}(y,z;t) \leq \varepsilon^{-2}$ and that μ_t satisfies the continuity equation

$$\partial_t \mu_t + \operatorname{div} \widetilde{Q}^\mu_t = 0. \tag{9.3}$$

The symbol $\widetilde{Q}^\mu$ in (9.3) is defined like in Definition 7.1 by $\widetilde{Q}^\mu_t(y,z) := \mu_t(y)\widetilde{r}(y,z;t)$. Trivially, $\widetilde{Q}^\mu = Q$ and therefore (9.3) follows from the definition of Λ_0. Due to Proposition 7.5, we conclude that $(\mu_t)_{t\geq 0}$ are the marginal distributions of the oscillatory steady state of the time inhomogeneous Markov chain with T_0-periodic jump rates (9.2).

We apply Lemma 9.1 considering the sequence $P_n := \mathbb{P}_x \circ (\mu^{(n)}, Q^{(n)})^{-1}$ and $\widetilde{P}^{(\mu,Q)}_n := \widetilde{\mathbb{P}}_x \circ (\mu^{(n)}, Q^{(n)})^{-1}$. The convergence $\widetilde{P}^{(\mu,Q)}_n \to \delta_{(\mu,Q)}$ follows by Lemma 7.3 and the above observation that μ_t is the marginal of the oscillatory steady state of $\widetilde{\mathbb{P}}$.

We now observe that

$$\frac{1}{n} \operatorname{Ent}\left(\widetilde{P}^{(\mu,Q)}_n \middle| P_n\right) \leq \frac{1}{n} \operatorname{Ent}\left(\widetilde{\mathbb{P}}_x|_{[0,nT_0]} \middle| \mathbb{P}_x|_{[0,nT_0]}\right). \tag{9.4}$$

This is a special case of a general result that says that relative entropy is decreasing under push forward. This follows directly by the variational representation of the relative entropy (see, e.g., [41, Sec. 8, Appendix 1]). By a direct computation, using (7.3), we have that the right hand side of (9.4) is given by

$$\widetilde{\mathbb{E}}_x\left[\mu^{(n)}\left(r - \widetilde{r}\right) + Q^{(n)}\left(\log \frac{\widetilde{r}}{r}\right)\right]. \tag{9.5}$$

Due to the definition of Λ_0 and by Assumption (A3), the functions $r - \widetilde{r}$ and $\log(\widetilde{r}/r)$ are bounded in modulus. Hence, by the $L^1(\mathbb{P}_x)$–convergence in (7.6) and (7.7) in Proposition 7.3, we get that, in the limit $n \to +\infty$, (9.5) converges to

$$\int_0^{T_0} I_t(\mu_t, Q_t)\mathrm{d}t = I(\mu, Q) = J(\mu, Q).$$

This completes the proof of (9.1).

It remains to prove that $(\mathrm{sc}^- J) = I$. Since I is lower semicontinuous and $I \le J$, by definition we have $(\mathrm{sc}^- J) \ge I$. We need to prove the converse inequality. Consider $(\mu, Q) \in \Lambda$. We construct a sequence $(\mu_n, Q_n) \in \Lambda_0$ such that $(\mu_n, Q_n) \to (\mu, Q)$ and moreover $\overline{\lim}_{n\to+\infty} J(\mu_n, Q_n) \le I(\mu, Q)$. This implies $(\mathrm{sc}^- J)(\mu, Q) \le I(\mu, Q)$ and allows to conclude the proof.

We construct the above sequence by a diagonal procedure. To this aim, we let $\Lambda_1 \subset \Lambda$ be the collection of elements $(\mu, Q) \in \Lambda$ such that there exists $\varepsilon > 0$ for which $\mu_t(x) \ge \varepsilon$ and $Q_t(y, z) \ge \varepsilon$ for all $t \in \mathcal{S}_{T_0}$, $x \in V$ and $(y, z) \in E$. Below we prove the following claim:

Claim 9.2. *The following holds:*

(i) *For any $(\mu, Q) \in \Lambda$, there exists a sequence $(\mu_n, Q_n) \in \Lambda_1$ such that $(\mu_n, Q_n) \to (\mu, Q)$ and moreover $\overline{\lim}_{n\to+\infty} I(\mu_n, Q_n) \le I(\mu, Q)$.*

(ii) *For any $(\mu, Q) \in \Lambda_1$, there exists a sequence $(\mu_n, Q_n) \in \Lambda_0$ such that $(\mu_n, Q_n) \to (\mu, Q)$ and moreover $\overline{\lim}_{n\to+\infty} I(\mu_n, Q_n) \le I(\mu, Q)$.*

The above claim allows to conclude as follows. Let us write $d(\cdot, \cdot)$ for a metric on $\mathcal{M}_*$ leading to the weak topology on $\mathcal{M}_*$ (see (3.11)). Fixed $(\mu, Q) \in \Lambda$, by Item (i) of the above claim, we can find (μ_n, Q_n) in Λ_1 with $d\left((\mu_n, Q_n), (\mu, Q)\right) \le 1/n$ and $I(\mu_n, Q_n) \le I(\mu, Q) + n^{-1}$. By Item (ii) we can find (μ_n^*, Q_n^*) in Λ_0 with $d\left((\mu_n^*, Q_n^*), (\mu_n, Q_n)\right) \le 1/n$ and $I(\mu_n^*, Q_n^*) \le I(\mu_n, Q_n) + n^{-1}$. Then $(\mu_n^*, Q_n^*) \to (\mu, Q)$ and $\overline{\lim}_{n\to+\infty} I(\mu_n^*, Q_n^*) \le I(\mu, Q)$. Using that $I(\mu_n^*, Q_n^*) = J(\mu_n^*, Q_n^*)$ by definition of J, we conclude that $\overline{\lim}_{n\to+\infty} J(\mu_n^*, Q_n^*) \le I(\mu, Q)$.

The rest of this section is devoted to the proof of Claim 9.2.

9.1. Proof of Item (i) in Claim 9.2

Let $(\mu, Q) \in \Lambda$. The sequence (μ_n, Q_n) is defined as

$$(\mu_n, Q_n) := \frac{1}{n}(\pi, Q^\pi) + \left(1 - \frac{1}{n}\right)(\mu, Q).$$

We point out that $\pi_t(y)$ can be estimated from below by the probability that $\xi_0 = y$ and that the Markov chain does not jump in the time interval $[0, t]$. Hence,

$$\pi_t(y) \ge \pi_0(y) \exp\left\{-\int_0^t r(y; s)\mathrm{d}s\right\}.$$

Due to Assumption (A3) (cf. (2.3)) and since π_0 is a positive measure, we conclude that $\min_y \inf_{t\in[0,T_0]} \pi_t(y) > 0$. As a by-product of this bound and again (2.3), we also conclude that $Q_t^\pi(y, z) = \pi_t(y) r(y, z; t)$ is bounded from below by a positive constant uniformly in $(y, z) \in E$ and $t \in [0, T_0]$. These observations imply that $(\pi, Q^\pi) \in \Lambda_1$ and therefore that $(\mu_n, Q_n) \in \Lambda_1$. Trivially, $(\mu_n, Q_n) \to (\mu, Q)$ in $\mathcal{M}_*$.

Since I is convex (cf. Remark 8.5) and $I(\pi, Q^\pi) = 0$, we have

$$I(\mu_n, Q_n) \le \left(1 - \frac{1}{n}\right) I(\mu, Q),$$

which implies that $\overline{\lim}_{n\to+\infty} I(\mu_n, Q_n) \le I(\mu, Q)$.

9.2. Proof of Item (ii) in Claim 9.2

Let $(\mu, Q) \in \Lambda_1$ and let $\varepsilon > 0$ be such that $\mu_t(x) \geq \varepsilon$ and $Q_t(y,z) \geq \varepsilon$ for all $t \in \mathcal{S}_{T_0}$, $x \in V$ and $(y,z) \in E$. To build (μ_n, Q_n) we fix a sequence of nonnegative C^∞–mollifiers φ_n with $\int \varphi_n(s)\mathrm{d}s = 1$ and with support in $[-1/n, 1/n]$ (see [9]). We write (μ_n, Q_n) for the element in $\mathcal{M}_*$ (cf. (3.11)) such that $\mu_n = [\mu_n]_t \mathrm{d}t$, $Q_n = [Q_n]_t \mathrm{d}t$ and

$$[\mu_n]_t(y) := \int_{\mathbb{R}} \mu_s(y)\varphi_n(t-s)\mathrm{d}s, \tag{9.6}$$

$$[Q_n]_t(y,z) := \int_{\mathbb{R}} Q_s(y,z)\varphi_n(t-s)\mathrm{d}s. \tag{9.7}$$

Since the maps $t \mapsto \mu_t(y)$ and $t \mapsto Q_t(y,z)$ are in $L^1(\mathrm{d}t) := L^1(\mathcal{S}_{T_0}, \mathrm{d}t)$, we have that the mollified maps $t \mapsto [\mu_n]_t(y)$ and $t \mapsto [Q_n]_t(y,z)$ are C^∞ and moreover converge in $L^1(\mathrm{d}t)$, as $n \to \infty$, to $t \mapsto \mu_t(y)$ and $t \mapsto Q_t(y,z)$, respectively (see, e.g., [9, Chp. 4]).

Let us first prove that $(\mu_n, Q_n) \in \Lambda_0$. It is simple to check that $(\mu_n, Q_n) \in \Lambda_1$ since $(\mu, Q) \in \Lambda_1$. Note in particular that they solve the continuity equation and that $[\mu_n]_t(y) \geq \varepsilon$ and $[Q_n]_t(y,z) \geq \varepsilon$. On the other hand, as already observed, $[\mu_n]_t$ and $[Q_n]_t$ depend smoothly on t; hence, they are bounded from above. This conclude the proof that $(\mu_n, Q_n) \in \Lambda_0$.

It remains to prove that $\overline{\lim}_{n\to\infty} I(\mu_n, Q_n) \leq I(\mu, Q)$. Since $(\mu, Q) \in \Lambda_1$ and due to Assumption (A3), we have that the following t functions

$$Q_t(y,z), \qquad Q_t(y,z)\log\mu_t(y), \qquad Q_t(y,z)\log r_t(y,z), \qquad \mu_t(y)r(y,z;t) \tag{9.8}$$

belong to $L^1(\mathrm{d}t)$. Hence, $I(\mu, Q)$ can be written as the sum among $(y,z) \in E$ of the following (y,z) parameterized expressions (which are meaningful since all terms below, with exception of at most one, are finite):

$$\begin{aligned} &\int_{\mathcal{S}_{T_0}} Q_t(y,z)\log Q_t(y,z)\mathrm{d}t - \int_{\mathcal{S}_{T_0}} Q_t(y,z)\log\mu_t(y)\mathrm{d}t \\ &- \int_{\mathcal{S}_{T_0}} Q_t(y,z)\log r(y,z;t)\mathrm{d}t - \int_{\mathcal{S}_{T_0}} Q_t(y,z)\mathrm{d}t + \int_{\mathcal{S}_{T_0}} \mu_t(y)r(y,z;t)\mathrm{d}t. \end{aligned} \tag{9.9}$$

Since the map $(0, +\infty) \ni u \mapsto u\log u \in \mathbb{R}$ is convex and since the mollification is an average, we have

$$[Q_n]_t(y,z)\log[Q_n]_t(y,z) \leq \int_{\mathbb{R}} \varphi_n(s)Q_{t-s}(y,z)\log Q_{t-s}(y,z)\mathrm{d}s.$$

Hence,

$$\begin{aligned} \int_{\mathcal{S}_{T_0}} [Q_n]_t(y,z)\log[Q_n]_t(y,z)\mathrm{d}t &\leq \int_{\mathcal{S}_{T_0}} \mathrm{d}t \int_{\mathbb{R}} \varphi_n(s)Q_{t-s}(y,z)\log Q_{t-s}(y,z)\mathrm{d}s \\ &= \int_{\mathcal{S}_{T_0}} Q_t(y,z)\log Q_t(y,z)\mathrm{d}t. \end{aligned} \tag{9.10}$$

On the other hand, due to Assumption (A3) and the properties of mollifiers stated after (9.7), we have the following limits in $L^1(\mathrm{d}t)$:

$$|[Q_n]_t(y,z) - Q_t(y,z)| \to 0, \tag{9.11}$$

$$|[Q_n]_t(y,z) - Q_t(y,z)||\log r(y,z;t)| \to 0, \tag{9.12}$$

$$|[\mu_n]_t(y) - \mu_t(y)|r(y,z;t) \to 0. \tag{9.13}$$

Finally, we estimate

$$\begin{aligned} &\big|[Q_n]_t(y,z)\log[\mu_n]_t(y) - Q_t(y,z)\log\mu_t(y)\big| \\ &\leq \big|[Q_n]_t(y,z) - Q_t(y,z)\big| \cdot \big|\log[\mu_n]_t(y)\big| \\ &\quad + Q_t(y,z)\big|\log\mu_t(y) - \log[\mu_n]_t(y)\big| \\ &\leq |\log\varepsilon| \cdot \big|[Q_n]_t(y,z) - Q_t(y,z)\big| + Q_t(y,z)\big|\log\mu_t(y) - \log[\mu_n]_t(y)\big|. \end{aligned} \tag{9.14}$$

We already know that the first term in the r.h.s. goes to zero in $L^1(\mathrm{d}t)$ (cf. (9.11)). On the other hand, we can bound

$$|Q_t(y,z)\big|\log\mu_t(y) - \log[\mu_n]_t(y)\big| \leq 2Q_t(y,z)|\log\varepsilon| \in L_1(\mathrm{d}t) \tag{9.15}$$

Since $t \mapsto [\mu_n]_t(y)$ converges to $t \mapsto \mu_t(y)$ in $L^1(\mathrm{d}t)$, at cost to extract a subsequence we can suppose that the convergence is also Lebesgue almost everywhere. As a by-product with (9.15), by dominated convergence, we conclude that also the second term in the r.h.s. of (9.14) goes to zero in $L_1(\mathrm{d}t)$, thus implying the limit

$$\big|[Q_n]_t(y,z)\log[\mu_n]_t(y) - Q_t(y,z)\log\mu_t(y)\big| \to 0 \text{ in } L_1(\mathrm{d}t). \tag{9.16}$$

To conclude, we write $I(Q_n,\mu_n)$ as the sum among $(y,z) \in E$ of

$$\begin{aligned} &\int_{\mathcal{S}_{T_0}} [Q_n]_t(y,z)\log[Q_n]_t(y,z)\mathrm{d}t - \int_{\mathcal{S}_{T_0}} [Q_n]_t(y,z)\log[\mu_n]_t(y)\mathrm{d}t \\ &- \int_{\mathcal{S}_{T_0}} [Q_n]_t(y,z)\log r(y,z;t)\mathrm{d}t - \int_{\mathcal{S}_{T_0}} [Q_n]_t(y,z)\mathrm{d}t + \int_{\mathcal{S}_{T_0}} [\mu_n]_t(y)r(y,z;t)\mathrm{d}t. \end{aligned} \tag{9.17}$$

Note that all the above integrals are finite since $(\mu_n, Q_n) \in \Lambda_0$ and due to Assumption (A3). By (9.10) the limsup of the first addendum in (9.17) is bounded from above by the first addendum in (9.9), while by using, respectively, (9.16), (9.12), (9.11), (9.13), we get that the limits of the other addenda in (9.17) are given by the similar addenda in (9.9). This concludes the proof that $\overline{\lim}_{n\to\infty} I(\mu_n, Q_n) \leq I(\mu, Q)$.

10. Proof of Theorem 3

Recall the continuous map $\mathcal{J} : \mathcal{M}_+(E \times \mathcal{S}_{T_0}) \to \mathcal{M}_a(E_s \times \mathcal{S}_{T_0})$ defined as $\mathcal{J}(Q)(y,z,A) := Q(y,z,A) - Q(z,y,A)$, with the convention that $Q(y',z',A) = 0$ if $(y',z') \notin E$. Due to the discussion preceding Definition 3.9, it only remains to show that the function

$$\widehat{I}(\mu, J) := \inf_{Q:\mathcal{J}(Q)=J} I(\mu, Q)$$

is convex and equals the r.h.s. of (3.21), and to derive (3.23).

Convexity follows from the convexity of I and the affinity of $\mathcal{J}$. Let us prove that $\widehat{I}(\mu, J)$ equals the r.h.s. of (3.21). Trivially, if $J = \mathcal{J}(Q)$ with $(\mu, Q) \in \Lambda$, then $(\mu, J) \in \Lambda_a$. On the other hand, if $(\mu, J) \in \Lambda_a$ then $J = \mathcal{J}(Q)$ where $Q_t(y, z) := \max\{J_t(y, z), 0\}$, in particular $(\mu, Q) \in \Lambda$. Since $I \equiv +\infty$ on Λ^c, we conclude that $\widehat{I}(\mu, J) = +\infty$ if $(\mu, J) \notin \Lambda_a$, in agreement with the r.h.s. of (3.21). Hence, from now on we restrict to $(\mu, J) \in \Lambda_a$.

Given a current $J \in \mathcal{M}_a(E_s \times \mathcal{S}_{T_0})$, we can write it uniquely in its Jordan decomposition $J = J^+ - J^-$. We recall that $J^\pm$ are nonnegative measures in $\mathcal{M}_+(E_s \times \mathcal{S}_{T_0})$ with disjoint supports. The antisymmetry of J implies that

$$J^+(y, z, A) = J^-(z, y, A) \qquad \forall A \subset \mathcal{S}_{T_0} \text{ measurable.}$$

Since we restrict to $(\mu, J) \in \Lambda_a$, we have $J^+ = J_t^+ \mathrm{d}t$ and $J^- = J_t^- \mathrm{d}t$, where $J_t^+(y, z) := \max\{J_t^+(y, z), 0\}$ and $J_t^-(y, z) := -\min\{J_t^+(y, z), 0\}$. Note that by property (v) in Definition 3.9 of Λ_a, J^+ and J^- have support included in $E \times \mathcal{S}_{T_0}$.

All the flows $Q \in \mathcal{M}_+(E \times \mathcal{S}_{T_0})$ such that $\mathcal{J}(Q) = J$ can be characterized by the decomposition $Q = J^+ + S$, where S is an arbitrary element of $\mathcal{M}_+(E \times \mathcal{S}_{T_0})$ such that

$$\begin{cases} S(y, z, A) = S(z, y, A) & \text{if } (y, z) \in E \text{ and } (z, y) \in E, \\ S(y, z, A) = 0 & \text{if } (y, z) \in E \text{ and } (z, y) \notin E. \end{cases}$$

Definition 10.1. *We denote by $\mathcal{S} = \mathcal{S}(\mu)$ the space of measures $S \in \mathcal{M}_+(E \times \mathcal{S}_{T_0})$ such that $S = S_t dt$, $S_t \in \mathbb{R}_+^E$,*

$$\begin{cases} S_t(y, z) = S_t(z, y) & \textit{if } (y, z) \in E \textit{ and } (z, y) \in E, \\ S_t(y, z) = 0 & \textit{if } (y, z) \in E \textit{ and } (z, y) \notin E, \end{cases}$$

and, given $(y, z) \in E$, if $\mu_t(y) = 0$ then $S_t(y, z) = 0$ for a.e. $t \in \mathcal{S}_{T_0}$.

Recall that we restrict to $(\mu, J) \in \Lambda_a$. By the previous observations, the flows Q such that $(\mu, Q) \in \Lambda$ and $\mathcal{J}(Q) = J$ are characterized by the decomposition $Q = J^+ + S$, where $S \in \mathcal{S}$.

Due to the previous observations, we have

$$\begin{aligned} \widehat{I}(\mu, J) &= \inf_{S \in \mathcal{S}} I(\mu, J^+ + S) \\ &= \inf_{S \in \mathcal{S}} \sum_{(y,z) \in E} \int_0^{T_0} \Phi\big(J_t^+(y, z) + S_t(y, z), \mu_t(y) r_t(y, z)\big) \mathrm{d}t, \end{aligned} \tag{10.1}$$

where the infimum is among the symmetric elements S as above. Note that we have set $r_t(y, z) := r(y, z; t)$. To solve the variational problem (10.1), it is enough to minimize for each t and for each $(y, z) \in E$ the contribution in the r.h.s. of (10.1) of the terms associated with (y, z) and to (z, y) (if $(z, y) \in E$; otherwise, one restricts only to the term associated with (y, z)).

To this aim, given $(v, w) \in E$ we set

$$Q_t^{J,\mu}(v, w) := \frac{J_t(v, w) + \sqrt{J_t^2(v, w) + 4\mu_t(v)\mu_t(w) r_t(v, w) r_t(w, v)}}{2}. \tag{10.2}$$

Case 1. For $(y,z) \in E$ with $(z,y) \notin E$ we know that $S_t(y,z) = 0$ and $J_t^+(y,z) = J_t(y,z)$ (see Definition 3.9-(v)). Therefore, for all $t \in [0,T_0]$, we have

$$\Phi\big(J_t^+(y,z) + S_t(y,z), \mu_t(y) r_t(y,z)\big) = \Phi\big(Q_t^{J,\mu}(y,z), \mu_t(y) r_t(y,z)\big).$$

Case 2. Let us now take $(y,z) \in E$ such that $(z,y) \in E$. It is enough to minimize, for each $t \in [0,T_0]$, the contribution

$$\begin{aligned} &\Phi\big(J_t^+(y,z) + S_t(y,z), \mu_t(y) r_t(y,z)\big) \\ &\quad + \Phi\big(J_t^+(z,y) + S_t(z,y), \mu_t(z) r_t(z,y)\big), \end{aligned} \tag{10.3}$$

when varying the parameter $S_t(y,z) = S_t(z,y)$ in $\mathbb{R}_+$. We define

$$s_t := S_t(y,z) = S_t(z,y), \quad j_t^+ := J_t^+(y,z), \quad j_t^- := J_t^-(y,z) = J^+(z,y).$$

Subcase 2.a. Supposing $\mu_t(y) > 0$ and $\mu_t(z) > 0$, by definition of Φ we have to minimize (cf. (10.3))

$$\begin{aligned} \inf_{s_t \in \mathbb{R}_+} \Big\{ &\big(j_t^+ + s_t\big) \log \frac{\big(j_t^+ + s_t\big)}{\mu_t(y) r_t(y,z)} + \big(j_t^- + s_t\big) \log \frac{\big(j_t^- + s_t\big)}{\mu_t(z) r_t(z,y)} \\ &+ \mu_t(y) r_t(y,z) + \mu_t(z) r_t(z,y) - j_t^+ - j_t^- - 2 s_t \Big\}. \end{aligned} \tag{10.4}$$

By simple computations, one gets that the minimizer is given by

$$s_t = \frac{-\big(j_t^+ + j_t^-\big) + \sqrt{(j_t^+ - j_t^-)^2 + 4 \mu_t(y) \mu_t(z) r_t(y,z) r_t(z,y)}}{2}.$$

We point out that $s_t > 0$ since $\min(j_t^+, j_t^-) = 0$. It then follows that the infimum in (10.4) equals

$$\Phi\big(Q_t^{J,\mu}(y,z), \mu_t(y) r_t(y,z)\big) + \Phi\big(Q_t^{J,\mu}(z,y), \mu_t(z) r_t(z,y)\big). \tag{10.5}$$

Subcase 2.b. If $\mu_t(y) = 0$ and $\mu_t(z) > 0$, then by Property (iv) in Definition 3.9 and by Definition 10.1 of $\mathcal{S}$ for a.e. t we have $j_t^+ = 0 = s_t$. In this case, for a.e. t the contribution (10.3) equals

$$j_t^- \log \frac{j_t^-}{\mu_t(z) r_t(z,y)} + \mu_t(z) r_t(z,y) - j_t^-, \tag{10.6}$$

which again equals (10.5).

Subcases 2.c, 2.d. If $\mu_t(y) > 0$ and $\mu_t(z) = 0$, or $\mu_t(y) = 0$ and $\mu_t(z) = 0$, one gets that $s_t = 0$ and the contribution (10.3) equals (10.5) by the same arguments used in Subcase 2.b.

Collecting all the above cases from Case 1 to Case 2.d, we get that

$$\widehat{I}(\mu, J) = \int_0^{T_0} I_t(\mu_t, Q_t^{J,\mu})\, \mathrm{d}t \tag{10.7}$$

for any $(\mu, J) \in \Lambda_a$. This concludes the proof of (3.21).

Finally, the derivation of (3.23) from the above formula can be done as in [6] (cf. Theorem 6.1 there) by adapting the conclusion there. Let us give more comments. Take $(\mu, J) \in \Lambda_a$. As for [6, Eq. (6.6)] we have

$$\Psi\left(J_t(y,z), J_t^\mu(y,z); a_t^\mu(y,z)\right) = \Phi\left(Q_t^{J,\mu}(y,z), \mu_t(y)r(y,z;t)\right) + \Phi\left(Q_t^{J,\mu}(z,y), \mu_t(z)r(z,y;t)\right)$$

if both (y,z) and (z,y) belong to E. Hence, in this case we have

$$\frac{1}{2}\left\{\Psi\left(J_t(y,z), J_t^\mu(y,z); a_t^\mu(y,z)\right) + \Psi\left(J_t(z,y), J_t^\mu(z,y); a_t^\mu(z,y)\right)\right\} = \Phi\left(Q_t^{J,\mu}(y,z), \mu_t(y)r(y,z;t)\right) + \Phi\left(Q_t^{J,\mu}(z,y), \mu_t(z)r(z,y;t)\right). \quad (10.8)$$

Let us now suppose that $(y,z) \in E$ and $(z,y) \notin E$. Then, it must be $J_t(y,z) = Q_t^{J,\mu}(y,z) \geq 0$ and $J_t^\mu(y,z) = \mu(y)r_t(y,z) \geq 0$. Since $a_t^\mu(y,z) = 0$, we have

$$\Psi\left(J_t(y,z), J_t^\mu(y,z); a_t^\mu(y,z)\right) = \Phi\left(Q_t^{J,\mu}(y,z), \mu_t(y)r(y,z;t)\right).$$

On the other hand, we have $a_t^\mu(z,y) = 0$, $J_t(z,y) = -J_t(y,z) \leq 0$ and $J_t^\mu(z,y) = -J_t^\mu(y,z) \leq 0$. Hence, by definition of Ψ, we have

$$\Psi\left(J_t(z,y), J_t^\mu(z,y); a_t^\mu(z,y)\right) = \Psi\left(J_t(y,z), J_t^\mu(y,z); a_t^\mu(y,z)\right).$$

Since moreover $\Phi\big(Q_t^{J,\mu}(z,y), \mu_t(z)r(z,y;t)\big) = \Phi(0,0) = 0$, also in this case we have (10.8). By symmetry we conclude that (10.8) holds for any $(y,z) \in E_s$. As a by-product of the above observation, (3.2) and (10.7), we get (3.23).

We conclude by discussing goodness and convexity of $\widehat{I}$. Goodness follows from the goodness of I by application of the contraction principle. On the other hand, by (3.19), $\widehat{I}(\mu, Q)$ equals the infimum of the convex rate functional I on a suitable affine subspace, thus implying that $\widehat{I}$ itself is convex.

11. Proof of Theorem 4

In what follows, as done before, we use the convention $0 \log 0 := 0$.

11.1. Proof of (5.2)

Since both (μ, Q) and $(\theta\mu, \theta Q)$ belong to Λ, we can write

$$\begin{aligned}
&I(\theta\mu, \theta Q; r) - I(\mu, Q; r) \\
&\quad = \sum_{y,z} \int_0^{T_0} \mathrm{d}s \left[\begin{array}{c} -Q_s(y,z) \log \frac{Q_s(y,z)}{\mu_s(y) r(y,z;s)} \\ +Q_{T_0-s}(z,y) \log \frac{Q_{T_0-s}(z,y)}{\mu_{T_0-s}(y) r(y,z;s)} \end{array}\right] \\
&\qquad + \sum_{y} \int_0^{T_0} \mathrm{d}s \left[\begin{array}{c} \sum_z Q_s(y,z) - \mu_s(y) r(y;s) \\ -\sum_z Q_{T_0-s}(z,y) + \mu_{T_0-s}(y) r(y;s) \end{array}\right] \\
&\quad = \sum_{y,z} \int_0^{T_0} \mathrm{d}s \left[\begin{array}{c} -Q_s(y,z) \log \frac{Q_s(y,z)}{\mu_s(y) r(y,z;s)} \\ +Q_s(z,y) \log \frac{Q_s(z,y)}{\mu_s(y) r(y,z;T_0-s)} \end{array}\right] \\
&\qquad + \sum_{y} \int_0^{T_0} \mathrm{d}s \left[\begin{array}{c} \sum_z Q_s(y,z) - \mu_s(y) r(y;s) \\ -\sum_z Q_s(z,y) + \mu_s(y) r(y;T_0-s) \end{array}\right].
\end{aligned}$$

Therefore, we have

$$I(\theta\mu,\theta Q;r)-I(\mu,Q;r)$$
$$=\int_0^{T_0}\mathrm{d}s\Big[\sum_{y,z}Q_s(y,z)\log\frac{\mu_s(y)r(y,z;s)}{\mu_s(z)r(z,y;T_0-s)}+\sum_y\mu_s(y)\left(-r(y;s)+r(y;T_0-s)\right)\Big]. \tag{11.1}$$

On the other hand, we have

$$\int_0^{T_0}\mathrm{d}s\sum_{y,z}Q_s(y,z)\log\frac{\mu_s(y)}{\mu_s(z)}$$
$$=\int_0^{T_0}\mathrm{d}s\sum_y\log\big(\mu_s(y)\big)\sum_z\left(Q_s(y,z)-Q_s(z,y)\right).$$

Using now the continuity equation $\partial_s\mu_s(y)+\sum_z\left[Q_s(y,z)-Q_s(z,y)\right]=0$, we obtain

$$\int_0^{T_0}\mathrm{d}s\sum_{y,z}Q_s(y,z)\log\frac{\mu_s(y)}{\mu_s(z)}=-\sum_y\int_0^{T_0}\mathrm{d}s\log\big(\mu_s(y)\big)\partial_s\mu_s(y)=0. \tag{11.2}$$

As a by-product of (11.1) and (11.2) one gets (5.2).

11.2. Proof of (5.3)

Since $(\mu,Q)\in\Lambda$, we can write

$$I(\theta\mu,\theta Q;r^{\mathrm{R}})-I(\mu,Q;r)=\sum_{y,z}\int_0^{T_0}\mathrm{d}s\begin{bmatrix}-Q_s(y,z)\log\frac{Q_s(y,z)}{\mu_s(y)r(y,z;s)}\\+Q_{T_0-s}(z,y)\log\frac{Q_{T_0-s}(z,y)}{\mu_{T_0-s}(y)r(y,z;T_0-s)}\end{bmatrix}$$
$$+\sum_y\int_0^{T_0}\mathrm{d}s\begin{bmatrix}\sum_z Q_s(y,z)-\mu_s(y)r(y;s)\\-\sum_z Q_{T_0-s}(z,y)+\mu_{T_0-s}(y)r(y;T_0-s)\end{bmatrix}.$$

By a local change of variable $T_0-s\mapsto s$ the last expression in the r.h.s. is zero, while the first expression can be simplified. This leads to

$$I(\theta\mu,\theta Q;r^{\mathrm{R}})-I(\mu,Q;r)=\sum_{y,z}\int_0^{T_0}\mathrm{d}s\,Q_s(y,z)\log\frac{\mu_s(y)r(y,z;s)}{\mu_s(z)r(z,y;s)}.$$

By (11.2) we can write the above r.h.s. as $\sum_{y,z}\int_0^{T_0}\mathrm{d}s\,Q_s(y,z)\log\frac{r(y,z;s)}{r(z,y;s)}=S_{\mathrm{tot}}(Q;r)$.

11.3. Proof of (5.4)

Since $(\mu,Q)\in\Lambda$, we can write

$$I(\theta\mu,\theta Q;r^{\mathrm{DR}})-I(\mu,Q;r)$$
$$=\sum_{y,z}\int_0^{T_0}\mathrm{d}s\begin{bmatrix}-Q_s(y,z)\log\frac{Q_s(y,z)}{\mu_s(y)r(y,z;s)}\\+Q_{T_0-s}(z,y)\log\frac{Q_{T_0-s}(z,y)}{\mu_{T_0-s}(y)w^{-1}_{T_0-s}(y)r(z,y;T_0-s)w_{T_0-s}(z)}\end{bmatrix}$$
$$+\sum_y\int_0^{T_0}\mathrm{d}s\begin{bmatrix}\sum_z Q_s(y,z)-\mu_s(y)r(y;s)\\-\sum_z Q_{T_0-s}(z,y)+\mu_{T_0-s}(y)r(y;T_0-s)\end{bmatrix}$$

$$
\begin{aligned}
&= \sum_{y,z} \int_0^{T_0} \mathrm{d}s \left[\begin{array}{c} -Q_s(y,z) \log \frac{Q_s(y,z)}{\mu_s(y) r(y,z;s)} \\ + + Q_s(z,y) \log \frac{Q_s(z,y)}{\mu_s(y) w_s^{-1}(y) r(z,y;s) w_s(z)} \end{array} \right] \\
&\quad + \sum_y \int_0^{T_0} \mathrm{d}s \left[\begin{array}{c} \sum_z Q_s(y,z) - \mu_s(y) r(y;s) \\ - \sum_z Q_s(z,y) + \mu_s(y) r(y;s) \end{array} \right] \\
&= \int_0^{T_0} \mathrm{d}s \sum_{y,z} Q_s(y,z) \log \frac{\mu_s(y) w_s^{-1}(y)}{\mu_s(z) w_s^{-1}(z)} \\
&= \int_0^{T_0} \mathrm{d}s \sum_{y,z} Q_s(y,z) \log \frac{w_s(z)}{w_s(y)}.
\end{aligned}
$$

We point out that the second identity follows from a local chance of variable $s \mapsto T_0 - s$, while the forth identity follows from (11.2).

By using the continuity equation $\partial_s \mu_s(z) = \sum_y [Q_s(y,z) - Q_s(z,y)]$ and integrating by parts, we conclude the proof of (5.4) by observing that

$$
\begin{aligned}
&\int_0^{T_0} \mathrm{d}s \sum_{y,z} Q_s(y,z) \log \frac{w_s(z)}{w_s(y)} = \int_0^{T_0} \mathrm{d}s \sum_z \log\big(w_s(z)\big) \sum_y \big(Q_s(y,z) - Q_s(z,y)\big) \\
&= \sum_z \int_0^{T_0} \mathrm{d}s \log\big(w_s(z)\big) \partial_s \mu_s(z) = -\sum_z \int_0^{T_0} \mathrm{d}s \mu_s(z) \partial_s \log\big(w_s(z)\big) = S_{\mathrm{ex}}(\mu; r).
\end{aligned}
$$

11.4. Proof of (5.5), (5.6) and (5.7)

These last three identities follow by minimizing (5.2), (5.3), (5.4), respectively. One needs to observe that the map $(\mu, Q) \mapsto (\theta\mu, \theta Q)$ is a bijection on Λ and to use the identities $S_{\mathrm{naive}}(\theta\mu, \theta Q; r) = -S_{\mathrm{naive}}(\mu, Q; r)$, $S_{\mathrm{tot}}(\theta Q; r^{\mathrm{R}}) = -S_{\mathrm{tot}}(Q; r)$, $S_{\mathrm{ex}}(\theta\mu; r^{\mathrm{DR}}) = -S_{\mathrm{ex}}(\mu; r)$. For the last identity, we observe that the accompanying measure w_s^{DR} associated with the rates $r^{\mathrm{DR}}(\cdot, \cdot; s)$ equals $w_{T_0 - s}$.

Acknowledgements

We thank the anonymous referees for their careful reading and suggestions. A. Faggionato and D. Gabrielli thank the Institute Henri Poincaré for the kind hospitality and the support during the trimester "Stochastic Dynamics Out of Equilibrium," in which they have worked on the manuscript.

References

[1] Barato, A.C., Seifert, U.: Thermodynamic uncertainty relation for biomolecular processes. Phys. Rev. Lett. **114**, 158101 (2015)

[2] Billingsley, P.: Convergence of probability measures. In: Wiley Series in Probability and Statistics: Probability and Statistics, 2nd edn. Wiley, New York (1999)

[3] Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G., Landim, C.: Macroscopic fluctuation theory. Rev. Mod. Phys. **87**, 593–636 (2015)

[4] Bertini, L., Faggionato, A., Gabrielli, D.: From level 2.5 to level 2 large deviations for continuous time Markov chains. Markov Process. Relat. Fields **20**, 545–562 (2014)

[5] Bertini, L., Faggionato, A., Gabrielli, D.: Large deviations of the empirical flow for continuous time Markov chains. Ann. Inst. H. Poincaré Probab. Stat. **51**, 867–900 (2015)

[6] Bertini, L., Faggionato, A., Gabrielli, D.: Flows, currents, and cycles for Markov chains: large deviation asymptotics. Stoch. Proc. Appl. **125**, 2786–2819 (2015)

[7] Blickle, V., Bechinger, C.: Realization of a micrometre-sized stochastic heat engine. Nat. Phys. **8**, 143–146 (2012)

[8] Brandner, K., Saito, K., Seifert, U.: Thermodynamics of micro-and nano-systems driven by periodic temperature variations. Phys. Rev. X **5**, 031019 (2015)

[9] Brezis, H.: Funcional Analysis, Sobolev Spaces and Partial Differential Equations. Universitext, Springer, New York (2011)

[10] Chernyak, V., Chertkov, M., Jarzynski, C.: Path-integral analysis of fluctuation theorems for general Langevin processes. J. Stat. Mech. P08001 (2006)

[11] Chernyak, V.Y., Sinitsyn, N.A.: Pumping restriction theorem for stochastic networks. Phys. Rev. Lett. **101**, 160601 (2008)

[12] Chernyak, V.Y., Chertkov, M., Malinin, S.V., Teodorescu, R.: Non-equilibrium thermodynamics and topology of currents. J. Stat. Phys. **137**, 109–147 (2009)

[13] Chetrite, R., Barato, A.C.: A formal view on level 2.5 large deviations and fluctuation relations. J. Stat. Phys. **160**(5), 1154–1172 (2015)

[14] Chetrite, R., Gawedzki, K.: Fluctuation relations for diffusion processes. Commun. Math. Phys. **282**, 469–518 (2008)

[15] Crooks, G.E.: Non-equilibrium measurements of free energy differences for microscopically reversible Markovian systems. J. Stat. Phys. **90**, 1481–1487 (1998)

[16] Crooks, G.E.: Path-ensemble averages in systems driven far from equilibrium. Phys. Rev. E **61**, 2361–2366 (2000)

[17] Davis, M.H.A.: Markov Models and Optimization. Chapman & Hall, London (1993)

[18] Dembo, A., Zeitouni, O.: Large Deviation Techniques and Applications, 2nd edn. Springer, New York (1998)

[19] Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time. Comm. Pure Appl. Math. (I) **28**, 1–47 (1975); (II) **28**, 279–301 (1975); (III) **29**, 389–461 (1976); (IV) **36**, 183–212 (1983)

[20] Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time. Comm. Pure Appl. Math. **II**(28), 279–301 (1975)

[21] Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time. Comm. Pure Appl. Math. **III**(29), 389–461 (1976)

[22] Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time. Comm. Pure Appl. Math. (IV) **36**, 183–212 (1983)

[23] Eelkema, R., Pollard, M.M., Vicario, J., Katsonis, N., Ramon, B.S., Bastiaansen, C.W.M., Broer, D.J., Feringa, B.L.: Nanomotor rotates microscale objects. Nature **440**, 163 (2006)

[24] De la Fortelle, A.: Large deviation principle for Markov chains in continuous time. Prob. Inf. Transm. **37**, 120 (2001)

[25] Faggionato, A., Gabrielli, D., Ribezzi Crivellari, M.: Non-equilibrium thermodynamics of piecewise deterministic Markov processes. J. Stat. Phys. **137**, 259–304 (2009)

[26] Faggionato, A., Mathieu, P.: Linear response and Nyquist relation in periodically driven Markov processes (Forthcoming)

[27] Gallavotti, G., Cohen, E.G.D.: Dynamical ensembles in nonequilibrium statistical mechanics. Phys. Rev. Lett. **74**, 2694–2697 (1995)

[28] Gammaitoni, L., Hänggi, P., Jung, P., Marchesoni, F.: Stochastic resonance. Rev. Mod. Phys. **70**, 223–287 (1998)

[29] Gingrich, T., Horowitz, J., Perunov, N., England, J.: Dissipation bounds all steady-state current fluctuations. Phys. Rev. Lett. **116**(12), 120601 (2016)

[30] Gingrich, T., Rotskoff, G., Horowitz, J.: Inferring dissipation from current fluctuations. J. Phys. A: Math. Theor. **50**(18), 184004 (2017)

[31] Ge, H., Jiang, D.-Q., Qian, M.: Reversibility and entropy production of inhomogeneous Markov chains. J. Appl. Probab. **43**, 1028–1043 (2006)

[32] Harris, R.J., Schütz, J.M.: Fluctuation theorems for stochastic dynamics. J. Stat. Mech. P07020 (2007)

[33] Hanggi, P., Thomas, H.: Stochastic processes: time evolution, symmetries and linear response. Phys. Rep. **88**, 207–319 (1982)

[34] Hatano, T., Sasa, S.: Steady-state thermodynamics of Langevin systems. Phys. Rev. Lett. **86**, 3463–3466 (2001)

[35] Höpfner, R., Kutoyants, Y.: Estimating discontinuous periodic signals in a time inhomogeneous diffusion. Stat. Inference Stoch. Proc. **13**, 193–230 (2010)

[36] Jensen, L.H.: Large deviations of the asymmetric simple exclusion process in one dimension. Ph.D. Thesis, Courant Institute NYU (2000)

[37] Joubaud, R., Pavliotis, G.A., Stoltz, G.: Langevin dynamics with space-time periodic nonequilibrium forcing. J. Stat. Phys. **158**, 1–36 (2015)

[38] Kaiser, M., Jack, R.L., Zimmer, J.: Canonical structure and orthogonality of forces and currents in irreversible Markov chains. J. Stat. Phys. **170**, 1019–1050 (2018)

[39] Kesidis, G., Walrand, J.: Relative entropy between Markov transition rate matrices. IEEE Trans. Inf. Theory **39**, 10561057 (1993)

[40] Kusuoka, S., Kuwada, K., Tamura, Y.: Large deviation for stochastic line integrals as L^p-currents. Probab. Theory Relat. Fields **147**, 649–667 (2010)

[41] Kipnis, C., Landim, C.: Scaling limits of interacting particle systems. Springer, Berlin (1999)

[42] Lazarescu, A.: The physicist's companion to current fluctuations: one-dimensional bulk-driven lattice gases. J. Phys. A **48**, 503001 (2015)

[43] Li, Q., Fuks, G., Moulin, E., Maaloum, M., Rawiso, M., Kulic, I., Foy, J.T., Giuseppone, N.: Macroscopic contraction of a gel induced by the integrated motion of light-driven molecular motors. Nat. Nanotechnol. **10**, 161–165 (2015)

[44] Maes, C.: The fluctuation theorem as a Gibbs property. J. Stat. Phys. **95**, 367–392 (1999)

[45] Maes, C., Netocný, K., Thomas, S.R.: General no-go condition for stochastic pumping. J. Chem. Phys. **132**, 234116 (2010)

[46] Maes, C., Netocný, K., Wynants, B.: Steady state statistics of driven diffusions. Physica A **387**, 2675 (2008)

[47] Maes, C., Netocný, K.: The canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states. Europhys. Lett. **82**, 30003 (2008)

[48] Mariani, M.: A Γ-convergence approach to large deviations. Ann. Sc. Norm. Super. Pisa Cl. Sci **18**, 951–976 (2018)

[49] Martinez, I.A., Roldán, É., Dinis, L., Petrov, D., Parrondo, J.M.R., Rica, R.A.: Brownian Carnot engine. Nat. Phys. **12**, 67–70 (2016)

[50] McNamara, B., Wiesenfeld, K.: Theory of stochastic resonance. Phys. Rev. A **39**, 4854 (1989)

[51] Mörters, P., Peres, Y.: Brownian Motion. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge (2010)

[52] Norris, J.R.: Markov Chains. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge (1999)

[53] Oono, Y., Paniconi, M.: Steady state thermodynamics. Prog. Theor. Phys. Suppl. **130**, 29–44 (1998)

[54] Pietzonka, P., Barato, A.C., Seifert, U.: Universal bounds on current fluctuations. Phys. Rev. E **93**(5), 052145 (2016)

[55] Pietzonka, P., Barato, A.C., Seifert, U.: Affinity- and topology-dependent bound on current fluctuations. J. Phys. A: Math. Theor. **49** (34), 34LT01 (2016)

[56] Proesmans, K., Van den Broeck, C.: Onsager coefficients in periodically driven systems. Phys. Rev. Lett. **115**, 090601 (2015)

[57] Proesmans, K., Cleuren, B., Van den Broeck, C.: Linear stochastic thermodynamics for periodically driven systems. J. Stat. Mech. 023202 (2016)

[58] Rahav, S., Horowitz, J., Jarzynski, C.: Directed flow in nonadiabatic stochastic pumps. Phys. Rev. Lett. **101**, 140602 (2008)

[59] Ray, S., Barato, A.C.: Stochastic thermodynamics of periodically driven systems: fluctuation theorem for currents and unification of two classes. Phys. Rev. E **96**, 052120 (2018)

[60] Reimann, P.: Brownian motors: noisy transport far from equilibrium. Phys. Rep. **361**, 57–265 (2002)

[61] Renger, M.D.R.: Large deviations of specific empirical fluxes of independent Markov chains, with implications for Macroscopic Fluctuation Theory. Weierstrass Institute, Preprint 2375 (2017)

[62] Rotskoff, G.M.: Mapping current fluctuations of stochastic pumps to nonequilibrium steady states. Phys. Rev. E **95**, 030101 (2017)

[63] Ruelle, D.: Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics. J. Stat. Phys. **95**, 393–468 (1999)

[64] 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)

[65] Sekimoto, K.: Stochastic energetics. Lecture Notes in Physics, vol. 799. Springer, Berlin (2010)

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

[67] Singh, N.: Onsager-Machlup theory and work fluctuation theorem for a harmonically driven Brownian particle. J. Stat. Phys. **131**, 405–414 (2008)

[68] Singh, N., Wynants, B.: Dynamical fluctuations for periodically driven diffusions. J. Stat. Mech. P03007 (2010)

[69] Sinitsyn, N.A., Akimov, A., Chernyak, V.Y.: Supersymmetry and fluctuation relations for currents in closed networks. Phys. Rev. E **83**, 021107 (2011)

[70] Verley, G., Van den Broeck, C., Esposito, M.: Modulated two-level system: exact work statistics. Phys. Rev. E **88**, 032137 (2013)

[71] Izumida, Y., Okuda, K.: Onsager coefficients of a finite-time Carnot cycle. Phys. Rev. E **80**, 021121 (2009)

[72] Izumida, Y., Okuda, K.: Linear irreversible heat engines based on local equilibrium assumptions. New J. Phys. **17**, 085011 (2015)

Lorenzo Bertini and Alessandra Faggionato
Dipartimento di Matematica
Università di Roma 'La Sapienza'
P.le Aldo Moro 2
00185 Rome
Italy
e-mail: bertini@mat.uniroma1.it;
faggiona@mat.uniroma1.it

Raphael Chetrite
CNRS Laboratoire J.A. Dieudonné
Université Côte d'Azur
06108 Nice Cedex 2
France
e-mail: rchetrit@unice.fr

and

Department of Physics Graduate School of Science
Kyoto University
Kyoto 606-8502
Japan

Davide Gabrielli
DISIM Università dell'Aquila
Via Vetoio, Loc. Coppito
67100 L'Aquila
Italy
e-mail: gabriell@univaq.it

Communicated by Christian Maes.

Received: December 22, 2017.

Accepted: June 27, 2018.