LEVEL 2.5 LARGE DEVIATIONS FOR CONTINUOUS TIME MARKOV CHAINS WITH TIME PERIODIC RATES

L. BERTINI, R. CHETRITE, A. FAGGIONATO, AND D. 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.

Keywords: time periodic Markov chain, large deviation principle, empirical measure, flow and current, Gallavotti–Cohen duality relation.

AMS 2010 Subject Classification: 60F10, 60J27, 82C05.

1. Introduction

The celebrated papers by Donsker and Varadhan [15] are a fundamental 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 hierarchal structure and the large deviation 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. 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 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.

A relevant motivation for the analysis of large deviation 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 deviations behavior, see e.g. [3, 30]. 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 [33, 34]. From a purely probabilistic viewpoint, the level 2.5 has been firstly investigated in [27] 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 [10]. In the same setting, the full large deviation principle is

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.
proven in [5] while the analogous result for diffusion processes is obtained in [28]. A more general setting with time dependent empirical measure and flow is considered in [26, 16]. We also point out that recently some thermodynamic uncertainly relation [1] and some related universal bound on current fluctuations [41, 42] have been derived in [18, 19] by using the level 2.5 large deviation principle. Finally, we refer to [9] for a more detailed discussion about level 2.5.

We here consider an irreducible continuous time Markov chain on a finite state space with time periodic jump rates and prove the corresponding level 2.5 large deviation principle. Continuous time Markov chains with periodic rates are natural models for several phenomena in different fields, among the significant applications we mention the molecular motors [45]. Time periodic forcing is also at the basis of several artificial micro-engines and molecular pumps (cf. e.g. [7, 31, 36]) and is also the natural context of stochastic resonance phenomena. For a time periodic protocol, the Gallavotti-Cohen symmetry for the entropy production rate has been experimentally verified in [49] and theoretically analyzed in stochastic models in [44, 48, 50, 51].

The derivation of the joint large deviation principle for the empirical measure and flow is here carried out according to the following scheme. Let $T_0$ be the period of the jump rates. On the time window $[0, T_0]$, we introduce a time dependent empirical measure whose value at time $t \in [0, T_0]$ is defined by sampling the chain at the times $t + kT_0$, $k = 0, \ldots, n$ and taking the empirical average. The time dependent empirical flow is defined by an analogous procedure. By an exponential change of the measure, we then prove the large deviation principle for these time dependent observables. The level 2.5 large deviation principle is then obtained by projection, the corresponding rate functional, as in the case of homogenization problems, is given by a variational formula relative to the period $[0, T_0]$. As an application of the above result, after introducing several forms of entropy production, we obtain the associated large deviation principles by projection. Moreover, we derive some Gallavotti Cohen duality relations at the level 2.5 and show that, by projection, theses relations imply other Gallavotti Cohen duality relations for the entropy production rate and some of them are new. In addition, we discuss in detail the case of a 2-state model.

Outline of the paper. In Section 2 we describe our main assumptions on the continuous time Markov chains with time periodic rates. In Section 3 we introduce the empirical measure, flow, current and state our main large deviation principles (cf. Theorems 1, 2 and 4). In Section 4 we discuss three forms of entropy production and in Section 5 we state the associated Gallavotti–Cohen duality relations (cf. Theorem 1 Corollaries 5.2 and 5.3). In Section 6 we apply our general results to the case of continuous time Markov chains with time periodic rates and two states. The rest of the paper is devoted to the proof of our results. In particular, in Section 7 we collect some preliminary facts. In Section 8 we prove the upper bound for the LDP stated in Theorem 2 (cf. Eq. (3.13)) and the convexity of the LD rate functional, while in Section 9 we prove the lower bound (cf. Eq. (3.13)) and the goodness of the rate functional. Theorem 4 follows easily from Theorem 2 by contraction and therefore the proof is omitted. The proofs of Theorems 5 and Theorem 4 are given in Sections 10 and 11 respectively.
2. Continuous time Markov chain 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 $r(x, x; s) \equiv 0$.

Roughly, the dynamics is defined as follows. Starting from a state $x$, the Markov chain spends at $x$ a random time $\tau_1$ such that

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

where

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

Knowing that $\tau_1 = t_1$, at time $t_1$ the Markov chain jumps to a new state $x_1$ chosen randomly with probability $r(x, x_1; t_1)/r(x; t_1)$, afterwards it waits in $x_1$ a random time $\tau_2$ such that

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

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 [13]. Indeed, we can introduce the continuous variable $s \in [0, +\infty)$ and describe the state of the system at time $t$ by $(\xi_t, s_t)$ 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 $L$

$$
L f(x, s) = \partial_s f(x, s) + \sum_y r(x, y; s) \left[ f(y, s) - f(x, s) \right].
$$

Following [13], to have a well defined operator one needs that the jump rates $r(x, y; \cdot)$ are measurable, locally integrable nonnegative functions. Due to [13] 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 $t > 0$, $y \neq z$.

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 graph $(V, E)$ is oriented–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 $D$ has zero Lebesgue measure, where $D \subset S_{T_0}$ is the set of discontinuity points of the jump rates $S_{T_0} \ni t \mapsto r(y, z; t) \in [0, \infty)$, as $y, z$ vary in $V$. 
Assumption (A2) means that, given two distinct sites $y, z$ in $V$, there is a family of vertexes $x_0, x_1, \ldots, x_n$ such that $x_0 = y$, $x_n = z$ and $(x_i, x_{i+1}) \in E$ for all $i = 0, 1, \ldots, n - 1$.

We point out that assumption (A4) is used only to derive Lemma 5.72.

Trivially, the discrete time process $\tilde{\xi} = (\tilde{\xi}_n)_{n \geq 0}$, with $\tilde{\xi}_n := \xi_{nT_0}$, is a time homogeneous Markov chain. We write $\tilde{p}(y, z)$, $y, z \in V$, for its jump probabilities. Since $V$ is finite and $(V, E)$ is oriented-connected, $\tilde{\xi}$ admits a unique invariant distribution $\pi$, i.e. a unique probability measure $\pi_0$ on $V$ such that

$$\sum_{y \in V} \pi_0(x) \tilde{p}(x, y) = \sum_{y \in V} \pi_0(y) \tilde{p}(y, x) \quad \forall \; x \in V. \quad (2.4)$$

Note that the Markov chain $\tilde{\xi}$ starting with the invariant distribution $\pi_0$ is stationary, i.e. its law is invariant by time shifts (cf. Th.1.7.1 in [39]). As a byproduct of this fact, the Markov property fulfilled by $\xi$ and the $T_0$-periodicity of the jump rates, one easily gets that the Markov chain $\xi$ starting with initial distribution $\pi_0$ is $T_0$-stationary, i.e. its law is invariant by time translations along times $T_0, 2T_0, 3T_0, \ldots$. In particular, when $\xi$ starts with distribution $\pi_0$, the law $\pi_t$ of $\xi_t$ is a $T_0$-periodic function from $\mathbb{R}_+$ to the space $P(V)$ of probability measures on $V$. We point out that $\pi_0$ is indeed the only initial distribution for which the Markov chain $\xi$ starting at $\pi_0$ 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 [47], this state is called nonequilibrium oscillatory state, shortly NOS).

We write $S_{T_0}$ 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). In what follows, we set $\pi = (\pi_t)_{t \in S_{T_0}}$.

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

$$T_{y,z} = \{t^{(1)}_{y,z} < t^{(2)}_{y,z} < \cdots\}$$

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

$$T_y = \{t^{(1)}_y < t^{(2)}_y < \cdots\}$$

for the superposition $\cup_z T_{y,z}$. It is known that $T_y$ is a Poisson point process on $(0, \infty)$ with rate $\lambda(y) := \sum_z \lambda(y, z)$ (i.e. $T_y$ 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 $U_{y,z} = (U_{y,z}^{(k)})_{k \geq 1}$ uniformly distributed on $[0, 1]$. The random objects given by $U_{y,z}$, $T_{y,z}$ 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 $T_y \cap (t, + \infty)$ and let $k, z$ be such that $s = t^{(k)}_{y,z}$ (they are well defined a.s.). Then $s = t^{(k)}_{y,z}$ 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 again, afresh.
3. LARGE DEVIATION PRINCIPLES

3.1. Joint LDP for the empirical measure and flow.

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

$$\check{\mu}_T(X) = \frac{1}{T} \int_0^T \delta_{X_t} dt, \quad \check{Q}_T(X)(y, z) = \frac{1}{T} \sum_{t \in [0, T]: X_{t-} \neq X_t} 1((X_{t-}, X_t) = (y, z)).$$

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

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

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

$$I_t(\check{\mu}, \check{Q}) := \sum_{(y, z) \in E} \Phi(\check{Q}(y, z), \check{\mu}(y)r(y; z)). \quad (3.2)$$

**Theorem 1.** For each $x \in V$, as $T \to +\infty$ the family of probability measures $\{\mathbb{P}_x \circ (\mu_T, Q_T)^{-1}\}$ 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

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

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

$$\begin{cases} \partial_t \mu_t + \text{div} Q_t = 0, \\ \frac{1}{T_0} \int_0^{T_0} \mu_t dt = \check{\mu}, \\ \frac{1}{T_0} \int_0^{T_0} Q_t dt = \check{Q}. \end{cases} \quad (3.4)$$

Theorem 1 follows easily by contraction from Theorem 2 below, and therefore we omit the proof.

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

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

Below we will often use the convention that, given a function $A: E \to \mathbb{R}$, we set $A(y, z) := 0$ if $(y, z) \notin E$. For example, due to this convention, we can rewrite (3.5) as $\text{div} A(y) = \sum_z A(y, z) - \sum_z A(z, y)$. Finally, the above continuity equation...
\[ \partial_t \mu_t + \text{div} \, Q_t = 0 \text{ in } (3.3) \] is thought of in weak sense, i.e. (using the time $T_0$-periodicity)

\[ \int_0^{T_0} \sum_y \mu_s(y) \partial_s f(y, s) ds = \int_0^{T_0} \sum_y \text{div} \, Q_s(y) f(y, s) ds, \quad (3.6) \]

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

We finally observe that if $\text{div} \, \bar{Q} \neq 0$ then $I(\bar{\mu}, \bar{Q}) = +\infty$. Indeed by taking time average of the continuity equation in (3.4) and using that $t \mapsto \mu_t$ is $T_0$-periodic we deduce $\text{div} \, \bar{Q} = 0$.

**Remark 3.2.** By contraction one derives from Theorem 7.1 both a LDP for the empirical measure $\bar{\mu}_t$ and a LDP for the empirical flow $Q_T$.

**Remark 3.3.** By the goodness of the rate function $I$ (see Theorem 2) the infimum in (3.3) is achieved whenever (3.4) admits a solution. By using Proposition 7.5 below, we then deduce that $I(\bar{\mu}, \bar{Q}) = 0$ if and only if $\bar{\mu} = \frac{1}{T_0} \int_0^{T_0} \pi_t dt$ and $\bar{Q}(y, z) = \frac{1}{T_0} \int_0^{T_0} \pi_t(y) r(y, z; t) dt$, $(y, z) \in E$.

### 3.2. Joint LDP for the extended empirical measure and flow

We introduce the space $\mathcal{M}_{+,T_0}(V \times S_{T_0})$ as the family of nonnegative measures on $V \times S_{T_0}$ with total mass equal to $T_0$, and the space $\mathcal{M}_{+}(E \times S_{T_0})$ as the family of nonnegative measures on $E \times S_{T_0}$ 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 $V \times S_{T_0}$ and on $E \times S_{T_0}$ are automatically bounded). We will often use the trivial identifications $\mathcal{M}_{+}(V \times S_{T_0}) \sim \mathcal{M}_{+}(S_{T_0})^V$ and $\mathcal{M}_{+}(E \times S_{T_0}) \sim \mathcal{M}_{+}(S_{T_0})^E$, 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** $\mu^{(n)} \in \mathcal{M}_{+,T_0}(V \times S_{T_0})$ and the extended empirical flow $Q^{(n)} \in \mathcal{M}_{+}(E \times S_{T_0})$ defined by

\[
\mu^{(n)}(x, dt) = \mu^{(n)}_t(x) dt \quad \text{where} \quad \mu^{(n)}_t(x) := \frac{1}{n} \sum_{k=0}^{n-1} \delta_{X_{t+kT_0}(x)}, \quad (3.7)
\]

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

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

We can identify functions $f : V \times S_{T_0} \to \mathbb{R}$ 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 $f : E \times S_{T_0} \to \mathbb{R}$. By means of this identification,
Theorem 2. Given\( J\) measures on \( I\) rate functional be the symmetrization of \( E\) such that:

\[
\bar{I} \subset M
\]

We also point out that \( f, g\) where \( E\) strictly positive jump rates, we let \( Q\) \( f: V \times S_{T_0} \to \mathbb{R} \), \( g: E \times S_{T_0} \to \mathbb{R} \),

\[
\mu^{(n)}(f) = \frac{1}{n} \int_0^{nT_0} f(X_t, t) dt, \quad f: V \times S_{T_0} \to \mathbb{R}, \quad (3.9)
\]

\[
Q^{(n)}(g) = \frac{1}{n} \sum_{t \in [0,nT_0]: X_t \neq X_{t-}} g(X_{t-}, X_t, t), \quad g: E \times S_{T_0} \to \mathbb{R}, \quad (3.10)
\]

where \( f, g\) are bounded and measurable.

To simplify the notation from now on we set

\[
\mathcal{M}_* := \mathcal{M}_{+, T_0}(V \times S_{T_0}) \times \mathcal{M}_+(E \times S_{T_0}). \quad (3.11)
\]

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

(i) \( \mu = \mu dt \) with \( \mu_t(V) = 1 \) for almost every \( t \in S_{T_0} \);

(ii) \( Q = Q_t dt \);

(iii) \( \partial_t \mu_t + \text{div} Q_t = 0 \);

(iv) for almost every \( t \in S_{T_0} \) 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

\[
I(\mu, Q) := \int_0^{T_0} I_t(\mu_t, Q_t) dt + \infty \quad \text{if} \quad (\mu, Q) \in \Lambda, \quad (3.12)
\]

\[
\text{otherwise.}
\]

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

\[
\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}_x((\mu^{(n)}, Q^{(n)}) \in \mathcal{C}) \leq - \inf_{(\mu, Q) \in \mathcal{C}} I(\mu, Q), \quad (3.13)
\]

\[
\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}_x((\mu^{(n)}, Q^{(n)}) \in \mathcal{A}) \geq - \inf_{(\mu, Q) \in \mathcal{A}} I(\mu, Q). \quad (3.14)
\]

We also point out that \( I(\mu, Q) = 0 \) if and only if \( \mu = \pi_t dt \) and \( Q = Q_t dt \) with \( Q_t(y, z) = \pi_t(y) r(y, z; t), (y, z) \in E\).

Since \( \bar{\mu}_{nT_0}(\cdot) = \frac{1}{nT_0} \mu^{(n)}(\cdot, S_{T_0}) \) and \( \bar{Q}_{nT_0}(\cdot) = \frac{1}{nT_0} Q^{(n)}(\cdot, S_{T_0}) \), Theorem 1 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 \( E_\alpha := \{(y, z) \in V \times V : (y, z) \in E \text{ or } (z, y) \in E\} \) be the symmetrization of \( E \) in \( V \times V \). We denote by \( \mathbb{R}_{E_\alpha}^E \) the family of functions \( \tilde{J} : E_\alpha \to \mathbb{R} \) which are antisymmetric, i.e. \( \tilde{J}(y, z) = -\tilde{J}(z, y) \forall (y, z) \in E_\alpha \).

**Definition 3.6.** Given \( T > 0 \), to each path \( \pi \in D(\mathbb{R}_+; V) \) we associate the empirical current \( \bar{J}_T(\pi) \in \mathbb{R}_{E_\alpha}^E \) defined as

\[
\bar{J}_T(\pi)(y, z) = \frac{1}{T} \sum_{t \in [0,T]: X_{t-} \neq X_t} \left[ \mathbb{1}(X_{t-}, X_t) = (y, z) - \mathbb{1}(X_{t-}, X_t) = (z, y) \right] \quad (3.15)
\]

for any \( (y, z) \in E_\alpha \).
To introduce the extended empirical current we denote by $\mathcal{M}_a(E_\mathbb{R} \times \mathcal{S}_{T_0})$ the space of signed measures $J$ on $E_\mathbb{R} \times \mathcal{S}_{T_0}$ which are antisymmetric in $E_\mathbb{R}$ (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 $\mathcal{M}_+(E_\mathbb{R} \times \mathcal{S}_{T_0})$). $\mathcal{M}_a(E_\mathbb{R} \times \mathcal{S}_{T_0})$ 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 $E_\mathbb{R} \times \mathcal{S}_{T_0}$.

**Definition 3.7.** Given $T > 0$, to each path $X \in D(\mathbb{R}_+; V)$ we associate the extended empirical current $J^{(n)}(X) \in \mathcal{M}_a(E_\mathbb{R} \times \mathcal{S}_{T_0})$ defined as

$$J^{(n)}(y, z, B) = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{t \in B + kT_0; \ X_{t-} \neq X_t} \left[ I\left((X_{t-}, X_t) = (y, z)\right) - I\left((X_{t-}, X_t) = (z, y)\right) \right]$$

(3.16)

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

We introduce the continuous map

$$\mathcal{J} : \mathcal{M}_+(E \times \mathcal{S}_{T_0}) \to \mathcal{M}_a(E_\mathbb{R} \times \mathcal{S}_{T_0})$$

defined as

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

for any $(y, z) \in E_\mathbb{R}$ 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:

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

(3.17)

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 $\hat{I}(\mu, J)$ given by

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

(3.18)

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

**Definition 3.8.** The set $\Lambda_a$ is given by the pairs $(\mu, J) \in \mathcal{M}_{+T_0}(V \times \mathcal{S}_{T_0}) \times \mathcal{M}_a(E_\mathbb{R} \times \mathcal{S}_{T_0})$ such that

(i) $\mu = \mu dt$ with $\mu(V) = 1$ for almost every $t \in \mathcal{S}_{T_0}$;
(ii) $J = Jdt$;
(iii) $\partial_t \mu_t + \text{div} \, J_t = 0$ where $\text{div} \, J_t(y) = \sum_{(y, z) \in E} J_t(y, z)$;
(iv) for almost every $t \in \mathcal{S}_{T_0}$, it holds: $\mu_t(y) = 0 \Rightarrow J_t(y, z) \leq 0$ for all $(y, z) \in E_\mathbb{R}$;
(v) $J_t(y, z) \geq 0$ if $(y, z) \in E$ and $(z, y) \notin E$, while $J_t(y, z) \leq 0$ 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\( ^1 \)

\[
\Psi(u, \bar{u}; a) := \begin{cases} 
  u \left[ \arcsinh \frac{u}{a} - \arcsinh \frac{\bar{u}}{a} \right] & \text{if } a > 0, \\
  - \left[ \sqrt{a^2 + u^2} - \sqrt{a^2 + \bar{u}^2} \right] & \text{if } a = 0.
\end{cases}
\] (3.19)

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

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 \( \{\mathbb{P}_x \circ (\mu^{(n)}, J^{(n)})^{-1}\}_{n \geq 1} \) of probability measures on

\( \mathcal{M}_{+,\mathcal{T}_0}(V \times \mathcal{S}_0) \times \mathcal{M}_a(E_s \times \mathcal{S}_0) \)

satisfies a large deviation principle with speed \( n \) and good and convex rate functional \( \tilde{I} \) given by

\[
\tilde{I}(\mu, J) = \begin{cases} 
  \int_0^{\mathcal{T}_0} I_1(\mu_t, Q^{J;\mu}_t) \, dt & \text{if } (\mu, J) \in \Lambda_a, \\
  +\infty & \text{otherwise},
\end{cases}
\] (3.20)

where

\[
Q^{J;\mu}_t(y, z) = \frac{J_t(y, z) + \sqrt{J_t^2(y, z) + 4\mu_t(y)\mu_t(z)r(y, z; t)r(z, y; t)}}{2}.
\] (3.21)

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

\[
\tilde{I}(\mu, J) = \frac{1}{2} \sum_{(y, z) \in E_s} \int_0^{\mathcal{T}_0} \Psi(J_t(y, z), J_t^\mu(y, z); a_t^\mu(y, z)) \, dt,
\] (3.22)

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 Section 10.

**Remark 3.9.** By contraction, Theorem 3 implies a joint LDP for the empirical measure and current. The corresponding convex rate functional \( \tilde{I} : \mathbb{P}(V) \times \mathbb{R}^E_s \mapsto [0, +\infty] \) is given by

\[
\tilde{I}(\tilde{\mu}, J) = \inf_{(\mu, J)} \frac{1}{\mathcal{T}_0} \tilde{I}(\mu, J),
\] (3.23)

where the infimum is taken among all pairs \((\mu, J) \in \Lambda_a\) such that \( \frac{1}{\mathcal{T}_0} \int_0^{\mathcal{T}_0} \mu_t dt = \tilde{\mu} \) and \( \frac{1}{\mathcal{T}_0} \int_0^{\mathcal{T}_0} J_t dt = \tilde{J} \), where \( \mu = \mu_t dt \) and \( J = J_t dt \).

### 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 \([0, n\mathcal{T}_0]\) as the Radon-Nykodym derivative

\[
\sigma_{n\mathcal{T}_0} [X] = \log \frac{d\mathbb{P}}{d(\mathbb{P}_B \circ R_{n\mathcal{T}_0})} \bigg|_{[0, n\mathcal{T}_0]} ((X_s)_{s \in [0, n\mathcal{T}_0]}),
\] (4.1)

\( ^1 \)This formula corrects the one in Eq. (6.3) when \( E_s \neq E \).
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 $\mu_0$ and $\mathbb{P}^B|_{[0,nT_0]}$ 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 $\mu'_0$. 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 pathwise time reversal at time $nT_0$. Of course, definition $4.1$ restricts to the case that the Radon-Nykodym derivative appearing there is indeed 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. Subsection 4.0.1), the reversed protocol first used in $[11]$ (cf. Subsection 4.0.2) and the dual reversed protocol first considered for time inhomogeneous processes in $[11]$ (cf. Subsection 4.0.3). The fact that by playing with different choices of the backward process (i.e. the rates $r^B$ here) we find different physics quantities (excess heat, housekeeping heat, phase-space contraction,...) has been pointed out in $[8, 10]$ for diffusion processes and in $[21]$ pure jump Markov processes.

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

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

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

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

By using formula $7.1$ in Section 7.1 and the periodicity of the rates, we get

$$\sigma_{nT_0}[X] = \log \frac{\mu_0(X_0)}{\mu'_0(X_{nT_0})} + \prod_{s \in (0,nT_0): X_s \neq X_{s'}} \log \frac{r(\cdot, X_s - ; X_s; s)}{r^B(\cdot, X_s - ; X_s; T_0 - s)} - \int_0^{nT_0} ds \left[r(X_s; s) - r^B(X_s; T_0 - s)\right].$$

We point out that, since $V$ is finite, the boundary term $\text{b.t.} = \log \frac{\mu_0(X_0)}{\mu'_0(X_{nT_0})}$ 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.

4.0.1. Entropy flow from naive reversal. We take the identity reversal $r^B(y, z; t) := r(y, z; t)$ and we write $\sigma_{nT_0}^{\text{naive}}$ 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:

$$S_{\text{naive}}(\mu, Q) := \sum_{(y,z) \in E} \int Q(y, z, ds) \log \frac{r(y, z; s)}{r(z, y; T_0 - s)} - \sum_y \int \mu(y, ds)[r(y; s) - r(y, T_0 - s)].$$

Then the entropy flow fulfills the identity

$$\frac{1}{n} \sigma_{nT_0}^{\text{naive}} = \text{b.t.} + S_{\text{naive}}(\mu^{(n)}, Q^{(n)}).$$

(4.6)
4.0.2. Total entropy flow from reversed protocol. As rates $r^B$ we choose the reversed protocol, i.e. we take the rates

$$r^R(y, z; t) := r(y, z; T_0 - t).$$  \hfill (4.7)

The resulting entropy flow, usually called total entropy flow, will be denoted by $\sigma_{\text{tot}}^{\text{tot}}$.

Defining the functional $S_{\text{tot}}$ as

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

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

$$\frac{1}{n} \sigma_{\text{tot}}^{\text{tot}} = \frac{b.t.}{n} + S_{\text{tot}}(Q^{(n)}).$$  \hfill (4.9)

The above entropy production has been investigated in [43, 44] for time periodic processes.

4.0.3. Entropy flow in excess. Given $t$ we write $w_t$ for the accompanying distribution (cf. [22]), 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 $r^B$ we then choose the dual reversed protocol:

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

The resulting entropy flow, denoted by $\sigma_{\text{ex}}^{nT_0}$, and called excess entropy flow, is related to the excess heat discussed in [23, 40]. By the invariance of $w_t$, from (4.10) one easily gets the simplified expression

$$\sigma_{nT_0}^{\text{ex}}[X] = \text{b.t.} + \sum_{s \in \{0, nT_0\}; X_s \neq X_n} \log \frac{w_s(X_s)}{w_s(X_n)}. \quad \hfill (4.11)$$

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

$$\sigma_{nT_0}^{\text{ex}}[X] = \text{b.t.}' - \int_0^{nT_0} ds \left[ \partial_s (\log w_s) \right] (X_s), \quad \hfill (4.12)$$

which is the quantity considered in the time periodic set-up by Schuller et al. in [49]. By defining the functional

$$S_{\text{ex}}(\mu) := -\sum_y \int \mu(y, ds) \partial_s (\log w_s)(y)$$  \hfill (4.13)

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

$$\frac{1}{n} \sigma_{nT_0}^{\text{ex}} = \frac{b.t.}{n} + S_{\text{ex}}(\mu^{(n)}). \quad \hfill (4.14)$$
5. Gallavotti–Cohen duality relations

As in Section 4 we assume that \((y, z) \in E\) if and only if \((z, y) \in E\), i.e. \(E = E_\ast\).

We recall that, given \((\mu, Q) \in M_\ast\) (cf. (3.11)), it holds \(I(\mu, Q) = +\infty\) if \((\mu, Q) \not\in \Lambda\). 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

\[
\theta \mu_t := \mu_{T_0 - t}, \quad \theta Q_t(y,z) := Q_{T_0 - t}(z,y).
\]

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

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 Section 4.0.1, 4.0.2 and 4.0.3). 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

\[
\begin{align*}
I_{\text{naive}}(s; r) &= \inf \{ I(\mu, Q; r) : S_{\text{naive}}(\mu, Q; r) = s \}, \\
I_{\text{tot}}(s; r) &= \inf \{ I(\mu, Q; r) : S_{\text{tot}}(Q; r) = s \}, \\
I_{\text{ex}}(s; r) &= \inf \{ I(\mu, Q; r) : S_{\text{ex}}(\mu; r) = s \}.
\end{align*}
\]

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

\[
\begin{align*}
I(\theta \mu, \theta Q; r) &= I(\mu, Q; r) + S_{\text{naive}}(\mu, Q; r), \quad (5.2) \\
I(\theta \mu, \theta Q; r^R) &= I(\mu, Q; r) + S_{\text{tot}}(Q; r), \quad (5.3) \\
I(\theta \mu, \theta Q; r^{DR}) &= I(\mu, Q; r) + S_{\text{ex}}(\mu; r). \quad (5.4)
\end{align*}
\]

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

\[
\begin{align*}
I_{\text{naive}}(-s; r) &= I_{\text{naive}}(s; r) + s, \quad (5.5) \\
I_{\text{tot}}(-s; r^R) &= I_{\text{tot}}(s; r) + s, \quad (5.6) \\
I_{\text{ex}}(-s; r^{DR}) &= I_{\text{ex}}(s; r) + s. \quad (5.7)
\end{align*}
\]

The above duality relations are new with exception of (5.6) which appears also in [44, 50]. The proof of Theorem 4 is given in Section 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 [49] 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 $\theta J_t(y, z) = J_{T_0 - t}(z, y)$ and write $\hat{I}(\mu, J; r)$ for the LD rate functional $\hat{I}(\mu, J)$ of $(\mu^{(n)}, J^{(n)})$ with jump rates $r(\cdot, \cdot, \cdot)$. In particular, one derives the following corollary (cf. (3.18)):

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

$$
\hat{I}(\theta \mu, \theta J; r^R) - \hat{I}(\mu, J; r) = \frac{1}{2} \sum_{(y, z) \in E} \int_0^{T_0} J_t(y, z) \log \frac{r(y, z; t)}{r(z, y; t)} \, dt, \\
\hat{I}(\theta \mu, \theta J; r^{DR}) - \hat{I}(\mu, J; r) = S_{\text{ex}}(\mu; r). 
$$

The proof of the above result is simple, hence omitted. 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

$$
\hat{I}(\theta J; r^R) - \hat{I}(J; r) = \frac{1}{2} \sum_{(y, z) \in E} \int_0^{T_0} J_t(y, z) \log \frac{r(y, z; t)}{r(z, y; t)} \, dt, \\
\hat{I}(\theta \mu; r^{DR}) - \hat{I}(\mu; r) = S_{\text{ex}}(\mu; r). 
$$

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 (5.11) even in time homogeneous set-up.

Finally we point out that the LD rate functional $\hat{I}(\bar{\mu}, \bar{Q})$ (cf. (3.3)) and $\hat{I}(\bar{\mu}, \bar{J})$ (cf. (3.23)) 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 3.6).

### 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.

For example in [43] we have a quantum dot with one single active energy level periodically modulated that corresponds to a two state Markov chain with rates

$$
\begin{cases}
  r_t(0, 1) = \frac{\Gamma}{1 + \exp(x_t)}, \\
  r_t(1, 0) = \frac{\Gamma}{1 + \exp(x_t)},
\end{cases}
$$

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 [49] we have a single defect center in natural IIa-type diamond exited by a red and a green laser with time periodic intensity. The corresponding rates are

\[
\begin{align*}
    r_t(0, 1) &= a_0(1 + \gamma \sin(\frac{2\pi}{\tau_0} t)), \\
    r_t(1, 0) &= b_0.
\end{align*}
\]  

(6.2)

In [37] we have a two state model of stochastic resonance given by

\[
\begin{align*}
    r_t(0, 1) &= \exp(-k \cos(\frac{2\pi}{\tau_0} t)), \\
    r_t(1, 0) &= \exp(k \cos(\frac{2\pi}{\tau_0} t)).
\end{align*}
\]  

(6.3)

Finally in [51] it is discussed a piecewise constant and symmetric protocol

\[
\begin{align*}
    r_t(0, 1) &= \exp(-h_t), \\
    r_t(1, 0) &= \exp(h_t), \\
    h_t &= \begin{cases} 
        h_0 - a & \text{if } 0 \leq t \leq \alpha T_0, \\
        h_0 + a & \text{if } \alpha T \leq t \leq T_0.
    \end{cases}
\end{align*}
\]  

(6.4)

Let us now discuss some results concerning the general situation. We restrict to elements Q and J absolutely continuous w.r.t. t. 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 \(\partial_t \mu_t + J_t = 0\).

With this notation \((\mu, Q) \in \Lambda\) if and only if \(\mu_t \in [0, 1]\) and \(\partial_t \mu_t + Q_t \geq 0\). Moreover, given \((\mu, Q) \in \Lambda\) the large deviations rate functional of Theorem 2 becomes

\[
I(\mu, Q) = \int_0^{T_0} \left[ Q_t \log \frac{Q_t}{\mu_t r_t(0, 1)} + (\partial_t \mu_t + Q_t) \log \frac{(\partial_t \mu_t + Q_t)}{(1 - \mu_t)r_t(1, 0)} + \mu_t r_t(0, 1) + (1 - \mu_t)r_t(1, 0) - 2Q_t \right] dt.
\]  

(6.5)

In this case, one can compute explicitly the LD rate functional \(I(\mu) = \inf_Q I(\mu, Q)\) associated to the extended empirical measure \(\mu^{(n)}\). We have that \(I(\mu)\) coincides in this case with the joint LD functional for measure and current, i.e \(I(\mu) = \hat{I}(\mu, J)\). 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, indeed it is enough that the unoriented graph obtained from the transition graph by disregarding the orientation is a tree). The rate functional \(I(\mu)\) is therefore obtained as \(I(\mu) = I(\mu, Q(\mu, \partial_t \mu))\) where (cf. [5.21])

\[
Q_{t}(\mu, \partial_{t} \mu) = \frac{-\partial_t \mu_t + \sqrt{(\partial_t \mu_t)^2 + 4\mu_t(1 - \mu_t)r_t(0, 1)r_t(1, 0)}}{2}.
\]  

(6.6)

Note that, despite the 2.5 level LDP (cf. [6.5]), even disregarding the constraints \(I(\mu)\) does not coincide with a time integration of the level 2 rate functional for the chain with frozen time dependence of the rates, i.e. with the following integral

\[
\int_0^{T_0} I_{s}^{\text{frozen}}(\mu_s) ds = \int_0^{T_0} \left( \sqrt{\mu_s r_s(0, 1)} - \sqrt{(1 - \mu_s) r_s(1, 0)} \right)^2 ds.
\]  

(6.7)

The two expressions coincide only in the limit of slow evolution (in particular, the zero order term of the formal expansion in \(\partial_t \mu\) of \(I(\mu, Q(\mu, \partial_t \mu))\) according to (6.0) coincides with (6.7)). Formula (6.7) follows by the explicit form of the level 2 rate functional for a 2–states chain, which is always reversible [15].
In [51] the large deviation 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 can be obtained by the contraction from our previous formulas.

The 2-states case is simple enough to allow also an explicit computation of the non-equilibrium oscillatory state \( \pi \). By a direct computation we have

\[
\pi_t(0) = \frac{e^{-\Gamma_t}}{1 - e^{-\Gamma_0}} \left[ \int_0^t r_s(1,0)e^{\Gamma_s} \, ds + e^{-\Gamma_0} \int_t^0 r_s(1,0)e^{\Gamma_s} \, ds \right]
\]

(6.8)

where \( \Gamma_t := \int_0^t [r_s(0,1) + r_s(1,0)] \, ds \). Note that this formula coincides with the one provided by [17, Prop. 3.13]. Recall that \( I(\mu, Q) \) is zero when \( \mu(y) = \pi_t(y) \) 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 \( \bar{I}(\bar{Q}) \) of the empirical flow \( \bar{Q}_T \) when \( T \to +\infty \) (see Remark [32]). By the graphical construction, since the jump rates are the same, we have that \( \bar{Q}_T \) coincides up to negligible terms with \( \frac{\bar{Q}_T}{\bar{T}} \) where \( \bar{N}_T \) is a non homogeneous Poisson process with periodic intensity given by \( r_t \). When \( T = n\bar{T}_0 \) we can write \( \bar{N}_T = \sum_{i=1}^n Y_i \), where the \( Y_i \) are i.i.d Poisson random variables of parameter \( \int_0^{\bar{T}_0} r_t \, dt \). The variable \( Y_i \) represents the number of points in the interval \( (i-1)\bar{T}_0, i\bar{T}_0 \]. Using the classic Cramer theorem we deduce that

\[
\bar{I}(\bar{Q}) = 2 \bar{Q}\log \left[ \frac{2\bar{Q}}{\bar{r}} \right] - 2\bar{Q} + \bar{r}, \quad \bar{r} := \frac{1}{\bar{T}_0} \int_0^{\bar{T}_0} r_t \, dt .
\]

(6.9)

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

\[
\bar{I}(\bar{Q}) := \frac{1}{\bar{T}_0} \inf_{(\mu, \bar{Q})} \left\{ \int_0^{\bar{T}_0} Q_t dt = \bar{Q} \right\} I(\mu, Q),
\]

(6.10)

is given by \( \mu_t = \frac{1}{T} \bar{Q} \) 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, we here obtain an upper bound for the rate function \( \bar{I} \) in term of the level 2.5 rate functional of a time homogenous Markov chain with suitable rates.

Let us call \( \bar{I}' \) the large deviations rate functional for the empirical measure and flow of a 2 states Markov chain having time independent rates equal to \( r(0,1) = r(1,0) = \bar{r} \). According to [3, 6] we have

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

(6.11)

where, by the divergence free condition, \( \bar{Q} := Q(0,1) = Q(1,0) \). By minimizing (6.11) among \( \bar{\mu} \) and comparing with (6.9) we get that

\[
\inf_{\bar{\mu}} \bar{I}'(\bar{\mu}, \bar{Q}) = \bar{I}(\bar{Q}) = \inf_{\bar{\mu}} \bar{I}(\bar{\mu}, \bar{Q}).
\]

In addition we can show the inequality

\[
\bar{I}(\bar{\mu}, \bar{Q}) \leq \bar{I}'(\bar{\mu}, \bar{Q}),
\]

(6.12)
which in general is strict. Inequality \((6.12)\) can be derived simply by inserting in \((5.3)\) the special pair \((\mu, \nu)\) given by

\[
\mu_t(y) = \bar{\mu}(y), \quad Q_t(y, z) = \frac{r_t(y, z)Q(y, z)}{\bar{r}}.
\]

Considering more general Markov chains one cannot expect inequality \((6.12)\) 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 8.3.

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 Section 2 the last assumption garantes that the associated continuous time Markov chain is well defined [13].

Given a probability measure \(\nu\) on \(V\), we write \(\mathbb{P}_\nu\) for the law of the Markov chain \((\xi_i)_{i \geq 0}\) with initial distribution \(\nu\), 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. Recall the definition of \(\pi = (\pi_t)_{t \in \mathcal{S}_{T_0}}\) given in Section 2.

**Definition 7.1.** Given \(\mu \in \mathcal{M}_{+}(0, \mathcal{S}_{T_0})\) 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–Nykodim 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 \(0 < t_1 < t_2 < \cdots < t_n < \tau_{N_t}\)

\[
\mathbb{P}_x(N_t = n, \tau_i \in [t_i, t_{i+1} + dt_i)\text{ for all }i = 1, 2, \ldots, n) = \exp\left\{-\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} r(x_i; s)ds \right\} \prod_{i=0}^{n-1} r(x_i, x_{i+1}; t_{i+1} - t_i)dt_1dt_2 \cdots dt_n,
\]

where \(x_0 := x, t_0 := 0\) and \(t_{n+1} := t\).

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

\[
\bar{r}(y, z; t) > 0 \quad \Rightarrow \quad r(y, z; t) > 0.
\]

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–Nykodim derivative on \(D([0, t]; V)\) is given by

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

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

\[
\bar{r}(y, z; t) = r(y, z; t)e^{F(y, z; t)}, \quad F(y, z; t) := \log \frac{\bar{r}(y, z; t)}{r(y, z; t)}.
\]

© 2019 L. Bertini, R. Chetrite, A. Faggionato, and D. Gabrielli
(above we used the convention \( \log(0/0) = 0 \)). Note that \( F \) is time periodic. Since \( r(y; \cdot) \) and \( r(y, z; \cdot) \) are \( T_0 \)-periodic functions, we can restate (7.1) as follows:

\[
\frac{d\tilde{\mathbb{P}}_x}{d\mathbb{P}_x}igg|_{[0,nT_0]} = \exp \left\{ n \mu^{(n)}(r - \tilde{r}) + nQ^{(n)}(F) \right\} , \quad \mathbb{P}_x := \tilde{\mathbb{P}}_x .
\] (7.2)

7.2. Some identities. Take \( Q \in M_+(E \times S_{T_0}) \). Denoting by \( \mathcal{B} \) the Borel sets of \( S_{T_0} \), for each \( y \in V \)

\[
\mathcal{B} \ni A \mapsto \sum_{z} Q(y, z, A) - \sum_{y} Q(z, y, A) =: \text{div} Q(y, A) \in \mathbb{R}
\]

is a signed measure on \( S_{T_0} \). In what follows we denote by \( \text{div} Q(f) \) the integral of \( f \) w.r.t. the above measure \( \text{div} Q \):

\[
\text{div} Q(f) = \sum_{y} \int_{0}^{T_0} \text{div} Q(y, ds) f(y, s) , \quad f : V \times S_{T_0} \to \mathbb{R} . \tag{7.3}
\]

**Lemma 7.2.** Let \( f : V \times S_{T_0} \to \mathbb{R} \) be \( C^1 \). Then

\[
\mu^{(n)}(\partial_s f) - \text{div} Q^{(n)}(f) = \frac{1}{n} (f(X_{nT_0},0) - f(X_0,0)) . \tag{7.4}
\]

**Proof.** Let \( s_1 < s_2 < \cdots < s_m \) 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

\[
f(X_{(k+1)T_0}, (k + 1)T_0) - f(X_{kT_0}, kT_0) = f(X_{s_m}, (k + 1)T_0) - f(X_{s_1}, kT_0) = \sum_{j=0}^{m} [f(X_{s_j}, s_{j+1}) - f(X_{s_j}, s_j)] + \sum_{j=1}^{m} [f(X_{s_j}, s_j) - f(X_{s_{j-1}}, s_j)]
\]

\[
= \sum_{j=0}^{m} \int_{s_j}^{s_{j+1}} \partial_s f(X_s, s) ds + \sum_{j=1}^{m} [f(X_{s_j}, s_j) - f(X_{s_{j-1}}, s_j)]
\]

\[
= \int_{kT_0}^{(k+1)T_0} \partial_s f(X_s, s) ds + \sum_{j=1}^{m} [f(X_{s_j}, s_j) - f(X_{s_{j-1}}, s_j)] .
\]

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

\[
\frac{1}{n} (f(X_{nT_0},0) - f(X_0,0)) = \mu^{(n)}(\partial_s f) + \sum_{y,z} \int_{[0,T_0]} Q^{(n)}(y, z, ds) (f(z, s) - f(y, s))
\]

\[
= \mu^{(n)}(\partial_s f) - \sum_{y} \sum_{z} \int_{[0,T_0]} Q^{(n)}(y, z, ds) f(y, s) + \sum_{y} \sum_{z} \int_{[0,T_0]} Q^{(n)}(z, y, ds) f(y, s)
\]

\[
= \mu^{(n)}(\partial_s f) - \text{div} Q^{(n)}(f) .
\]

\( \square \)

7.3. The oscillatory steady state. We collect in the following proposition some asymptotic properties of the oscillatory steady state. Recall 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 \( \pi_t \) as \( n \) goes to \( \infty \);

(ii) \( \mathbb{P}_x \)-a.s. \( \mu^{(n)} \) weakly converges to \( \pi_t dt \) in \( M_{+,T_0}(V \times S_{T_0}) \);
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 $\pi_t$ introduced in Section 7. This concludes the proof of Item (i). The proof of Items (ii), (iii), (iv) can be derived from [24, Theorem 2.1]. □

Lemma 7.4. It holds $\partial_t \pi_t + \text{div} Q^\mu_t = 0$.

Proof. Due to Definition 7.1 we only need to prove that $\mu^n(\partial_s f) - \text{div} Q^n(f) = 0$ for any $C^1$ function $f : V \times S_{T_0}$. 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 $\pi = \pi_t dt$.

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

\[ \partial_t \mu_t + \text{div} Q^\mu_t = 0 \]  

is given by $\pi$.

Proof. We first show that $\mu \in \mathcal{M}_{+,T_0}(V \times S_{T_0})$ with $\mu = \mu_t dt$ solving (7.6) is an invariant measure of the piecewise deterministic Markov process $(\xi_t, Y_t)_{t \geq 0}$ on $V \times S_{T_0}$, where $Y_t$ denotes the canonical projection of $t$ in $S_{T_0}$. By [13] this PDMP has extended generator $L$ 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 $V \times S_{T_0}$ via the canonical projection for times). By [13, Theorem (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 [13, Prop. 34.6], $\mu$ is an invariant measure for the PDMP if and only if $\mu(Lf) = 0$ 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 $\mu$ is a weak solution of (7.5).

Using now that $\mu$ is invariant for the PDMP, and therefore it equals the distribution at any $t \geq 0$, we get

\[ \sum_x \int_0^{T_0} \mu_s(x) f(x, s) ds = \sum_x \sum_y \int_0^{T_0} \mu_s(x)p_{s,s+t}(x, y)f(y, s+t) ds \]  

(7.6)

where $p_{s,s+t}(x, y) = P(\xi_{t+s} = y | \xi_s = x)$. We rewrite (7.6) as

\[ \sum_x \int_0^{T_0} \mu_s(x) f(x, s) ds = \sum_y \sum_{t} \int_{t}^{t+T_0} \mu_{w-t}(y)p_{w-t,w}(y, x)f(x, w) dw \]  

(7.7)

Since $\sum_x \int_0^{T_0} \mu_s(x) f(x, s) ds = \sum_x \int_{t}^{t+T_0} \mu_w(x)f(x, w) dw$ we get, for any $t \geq 0$, that

\[ \mu_w(x) = \sum_y \mu_{w-T_0}(y)p_{w-T_0,w}(y, x) \]

for almost all $w \in [t, t + T_0]$. Take $t = T_0$ and fix a $w \in [0, T_0]$ satisfying the above equation which now reads

\[ \mu_w(x) = \sum_y \mu_{w-T_0}(y)p_{w-T_0,w}(y, x) \]
Then $\mu_w$ is an invariant distribution of the Markov chain $(\xi_{w+nT_0})_{n \geq 0}$ and we know it is unique since $E$ is oriented connected. On the other hand, we know that $\pi_w$ has the same property, hence it must be $\mu_w = \pi_w$. We then conclude that $\mu_w = \pi_w$ for almost all $w \in [0, T_0]$, thus implying that $\mu_t dt = \pi_t dt$. Hence, $\mu = \pi$.

\begin{flushright}
\Box
\end{flushright}

8. Proof of Theorem 2: Upper bound (8.13) and convexity of $I$.

We start by showing exponential tightness:

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

**Proof.** Given $\ell > 0$ we set $K_\ell := \{(\mu, Q) \in \mathcal{M}_*: Q(1) \leq \ell\}$. Then $K_\ell$ is a compact subset of $\mathcal{M}_*$. To prove the exponential tightness it is enough to show that there exists $C > 0$ such that

$$
\lim_{n \to \infty} \frac{1}{n} \log P_x(\mu^{(n)}, Q^{(n)} \notin K_\ell) \leq -C\ell
$$

for large $\ell$.

We prove (8.1). The event $\{(\mu^{(n)}, Q^{(n)}) \notin K_\ell\}$ is simply the event that the measure $Q^{(n)}$ has total mass larger than $\ell$. Due to (8.2), 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 Section 2.1 the number of jumps in the time interval $[0, nT_0]$ is stochastically dominated by a Poisson variable $Z$ of parameter $\lambda nT_0$ where $\lambda = \sum_{(y, z)} \sup_{t \in [0, T_0]} r(y, z; t)$. Since $E[\exp(\gamma Z)] = \exp(\lambda nT_0(\exp(\gamma) - 1))$, by applying Chebyshev inequality we get

$$
P_x(\mu^{(n)}, Q^{(n)} \notin K_\ell) = P_x(Q^{(n)}(1) > \ell) \leq P(Z > n\ell) \leq e^{-n\ell} E[\exp(\gamma Z)] = \exp(-n\ell + \lambda nT_0(\exp(\gamma) - 1)).
$$

The above bound trivially implies (8.1).

\begin{flushright}
\Box
\end{flushright}

Given a $C^1$ function $\phi : V \times \mathcal{S}_{T_0} \to \mathbb{R}$ and given a continuous function $F : E \times \mathcal{S}_{T_0} \to \mathbb{R}$ we set

$$
r^F(y, z; t) = r(y, z; t)e^{F(y, z; t)},
$$

$$
r(y; t) = \sum_z r(y, z; t),
$$

$$
r^F(y; t) = \sum_z r^F(y, z; t),
$$

and we define $\hat{I}_{\phi, F} : \mathcal{M}_* \to \mathbb{R}$ and $I_{\phi, F} : \mathcal{M}_* \to [0, +\infty]$ as follows:

$$
\hat{I}_{\phi, F}(\mu, Q) := -\mu(\partial_1 \phi) + \text{div} Q(\phi) + Q(F) - \mu(r^F - r),
$$

$$
I_{\phi, F}(\mu, Q) := \begin{cases} 
\hat{I}_{\phi, F}(\mu, Q) & \text{if } \mu = \mu_t dt, \, \mu_t(V) = 1 \text{ a.s.} \\
+\infty & \text{otherwise.}
\end{cases}
$$

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

**Proof.** Let us call $\mathcal{A}$ be the set of pairs $(\mu, Q) \in \mathcal{M}_*$ such that $\mu = \mu_t dt$, $\mu_t(V) = 1$ a.s. It is simple to check that $\mathcal{A}$ is convex and closed in $\mathcal{M}_*$. Since $\mathcal{A}$ is convex and $\hat{I}_{\phi, F}$ is convex, it is simple to derive that $I_{\phi, F}(\mu, Q)$ is convex.
Lemma 8.3. Fix App. 2, we conclude that as a byproduct of the above bound and the minmax lemma (cf. [29, Lemma 3.3], \(\mu\) holds compact subsets \(K \subset M\) thus implying the thesis. □

Due to (7.4) we can write

\[ P \]

where \(P\) is the weak convergence of measures and since \(\partial_t \phi, \phi \) 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 \(h(y;t) = \sum r(y, z; t) e^{F(y, z; t)} - 1\), and \(F\) is regular in time, for each \(y\) the function \(h(y, \cdot)\) is continuous on \(S_{T_0} \setminus D\) (recall Assumption (A4)). On the other hand, since \(\nu = \nu(t) dt\), we have \(\sum y \nu(y, D) = 0\). As a byproduct of the last observation and the Portmanteau theorem as stated in [38, Thm. 12.6], we get that \(\nu^{(k)}(h) \to \nu(h)\). This proves that \(I_{\phi, F}^{\star}\) is continuous on the set \(A\). Since \(I_{\phi, F}\) is continuous on the closed set \(A\) and it equals \(+\infty\) on \(M_\ast \setminus A\), we conclude that \(I_{\phi, F}\) is lower semicontinuous.

We recall by (7.2) that

\[ \frac{dP^F_x}{dP_x}\bigg|_{[0, t_0]} = M^F_n \]

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

Due to (7.4) we can write

\[ -nI_{\phi, F}(\mu^{(n)}, Q^{(n)}) = \phi(X_{nT_0}, 0) - \phi(X_0, 0) - \log M^F_n. \]  

In the above identity we have used also that \(\mu^{(n)}(x, dt) = \mu^{(n)}(x) dt\) where \(0 \leq \mu^{(n)}(x) \leq 1\) (cf. [37]), thus implying that \(I_{\phi, F}(\mu^{(n)}, Q^{(n)}) = \tilde{I}_{\phi, F}(\mu^{(n)}, Q^{(n)})\).

Lemma 8.3. Fix \(x \in V\). For each \(\phi, F\) as above and each measurable \(B \subset M_\ast\) it holds

\[ \lim_{n \to \infty} \frac{1}{n} \log P_x \left( (\mu^{(n)}, Q^{(n)}) \in B \right) \leq - \inf_{(\mu, Q) \in B} I_{\phi, F}(\mu, Q). \]  

Proof. Due to (7.4) we can write

\[ P_x \left( (\mu^{(n)}, Q^{(n)}) \in B \right) = E_x \left( \exp \left\{ -nI_{\phi, F}(\mu^{(n)}, Q^{(n)}) - [\phi(X_{nT_0}, 0) - \phi(X_0, 0)] M^F_n 1_B(\mu^{(n)}, Q^{(n)}) \right\} \right) \]

\[ \leq \left[ \sup_{(\mu, Q) \in B} e^{-nI_{\phi, F}(\mu, Q)} \right] e^{2||\phi||_\infty} E_x(M^F_n) = \left[ \sup_{(\mu, Q) \in B} e^{-nI_{\phi, F}(\mu, Q)} \right] e^{2||\phi||_\infty}, \]

thus implying the thesis. □

Due to the exponential tightness, it is enough to prove the upper bound [M.1] for compact subsets \(K \subset M_\ast\) instead of generic closed subset \(C \subset M_\ast\). Due to Lemma 8.3 we have

\[ \lim_{n \to \infty} \frac{1}{n} \log P_x \left( (\mu^{(n)}, Q^{(n)}) \in K \right) \leq - \sup_{\phi, F, (\mu, Q) \in K} I_{\phi, F}(\mu, Q). \]

As a byproduct of the above bound and the minmax lemma (cf. [29, Lemma 3.3, App. 2]), we conclude that

\[ \lim_{n \to \infty} \frac{1}{n} \log P_x \left( (\mu^{(n)}, Q^{(n)}) \in K \right) \leq - \inf_{(\mu, Q) \in K} \sup_{\phi, F} I_{\phi, F}(\mu, Q). \]
Hence, to conclude the proof of the upper bound \((3.13)\) it is enough to apply the following lemma:

**Lemma 8.4.** For each \((\mu, Q) \in M_\star\) it holds

\[
I(\mu, Q) = \sup_{\phi,F} I_{\phi,F}(\mu, Q).
\]

where the supremum is taken among all \(C^1\) functions \(\phi : V \times S_{T_0} \rightarrow \mathbb{R}\) and continuous functions \(F : E \times S_{T_0} \rightarrow \mathbb{R}\).

**Remark 8.5.** Note that, by the convexity of \(I_{\phi,F}\), the above lemma implies the convexity of \(I\).

**Proof.** In what follows we write \(\ell(\cdot)\) for the Lebesgue measure on \(S_{T_0}\).

- **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 \(I_{\phi,F}(\mu, Q)\) one trivially gets that both sides of \((8.8)\) are \(+\infty\) if \((\mu, Q) \notin \Lambda\), where \(\Lambda\) is defined as in the proof of Lemma 8.2. It also trivial to verify that both sides of \((8.8)\) are \(+\infty\) if for some \(C^1\) function \(\varphi : V \times S_{T_0} \rightarrow \mathbb{R}\) it holds \(-\mu(\partial_y \varphi) + \text{div } Q(\varphi) \neq 0\). Hence, in what follows we restrict to the case \(\mu = \mu dt, \mu(V) = 1\) a.s., and \(\partial_y \mu + \text{div } Q = 0\) (in the weak sense). Since in this case \(I_{\phi,F}(\mu, Q)\) does not depend on \(\varphi\), we write simply \(I_F(\mu, Q)\).

Suppose now that \(Q\) is not of the form \(Q dt\). Hence there exists a subset \(B\) of \(S_{T_0}\) 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 \(D_\varepsilon\) and an open set \(G_\varepsilon\) 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 Urisohn lemma we can find a continuous function \(\varphi_\varepsilon : S_{T_0} \rightarrow [0,1]\) such that \(\varphi_\varepsilon \equiv 1\) on \(D_\varepsilon\) and \(\varphi_\varepsilon \equiv 0\) on \(G_\varepsilon^c\). 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

\[
I_{F_\varepsilon}(\mu, Q) = \sum_{y,z} \int_0^{T_0} Q(y, z, dt)F_\varepsilon(y, z, t) - \sum_y \int_0^{T_0} \mu_t(y)(rF_\varepsilon(y, t) - r(y, t)) dt
\]

\[
= \int_0^{T_0} Q(y_0, z_0, dt)F_\varepsilon(y_0, z_0, t) - \int_0^{T_0} \mu_t(y)r(y_0, z_0, t)e^{F_\varepsilon(y_0, z_0, t)} dt
\]

\[
\geq \gamma(\varepsilon)Q(y_0, z_0, D_\varepsilon) - e^{\gamma(\varepsilon)}\int_0^{T_0} \mu_t(y)r(y_0, z_0, t) dt
\]

\[
\geq \gamma(\varepsilon)Q(y_0, z_0, D_\varepsilon) - e^{\gamma(\varepsilon)}\ell(G_\varepsilon) \max_{y,z,t} r(y, z, t)
\]

\[
\geq \gamma(\varepsilon)Q(y_0, z_0, B)/2 - e^{\gamma(\varepsilon)}\max_{y,z,t} r(y, z, t).
\]

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 \(\Lambda\). Let us suppose that there exists \(B \subset S_{T_0}\) 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 \(D_\varepsilon\) and an open set \(G_\varepsilon\) such that \(D_\varepsilon \subset B \subset G_\varepsilon\).
and $\ell(G_x \setminus D_x) \leq \varepsilon$. Without loss we take $D_x \subset D_{x'}$, if $\varepsilon > \varepsilon'$. Since $\ell(B) > 0$ we have $\ell(D_{x'}) \geq \ell(B)/2 > 0$ for any $\varepsilon_0 := \ell(B)/2$. In particular, $\int_{D_{x'}} Q_t(y_0, z_0) \, dt > 0$. Hence, similarly to (8.9), we get

$$I_{F_t}(\mu, Q) \geq \gamma(\varepsilon) \int_{D_{x'}} Q_t(y_0, z_0) \, dt - e^{\gamma(\varepsilon)} \ell(G_x \setminus D_x) \max r(y, z, t).$$

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

This concludes the proof of our initial claim.

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

$$I(\mu, Q) = \sup_{\overline{F}} I_{F}(\mu, Q),$$

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

$$I_{F}(\mu, Q) = \sum_{(y, z)} \int_0^{T_0} Q_t(y, z) F(y, z, t) \, dt - \sum_y \int_0^{T_0} \mu_t(y)(r_F(y, t) - r(y, t)) \, dt$$

$$= \sum_{(y, z)} \int_0^{T_0} dt \left[ Q_t(y, z) F(y, z, t) - \mu_t(y) r(y, z, t)(e^{F(y, z, t)} - 1) \right].$$

(8.11)

Since (cf. (8.1)) $\Phi(q, p) = \sup_{v \in \mathbb{R}} \{qv - p(e^v - 1)\}$ 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 $\Phi(Q_t(y, z), \mu_t(y) r(y, z, t))$, thus implying that

$$I_{F}(\mu, Q) \leq \int_0^{T_0} \Phi(Q_t(y, z), \mu_t(y) r(y, z, t)) \, dt = I(\mu, Q).$$

(8.12)

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

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

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

$$A(y, z) := \{ t \in S_{T_0} : Q_t(y, z) = 0 \},$$

$$B(y, z) := \{ t \in S_{T_0} : \mu_t(y) = 0 \mathrm{~and~} Q_t(y, z) > 0 \},$$

$$C(y, z) := S_{T_0} \setminus (A(y, z) \cup B(y, z)) = \{ t \in S_{T_0} : Q_t(y, z) > 0 \mathrm{~and~} \mu_t(y) > 0 \},$$

$$C_{\varepsilon}(y, z) := \{ t \in S_{T_0} : \varepsilon \leq Q_t(y, z) \leq \frac{1}{\varepsilon} \mathrm{~and~} \varepsilon \leq \mu_t(y) \leq \frac{1}{\varepsilon} \}.$$

Since $(\mu, Q) \in \Lambda$, we have $\ell(B(y, z)) = 0$. In particular, by definition of $\Phi$ (cf. (5.1)),

$$\int_0^{T_0} \Phi(Q_t(y, z), \mu_t(y) r(y, z, t)) \, dt = \int_{A(y, z)} \mu_t(y) r(y, z, t) \, dt + \int_{C(y, z)} \Phi(Q_t(y, z), \mu_t(y) r(y, z, t)) \, dt.$$

(8.13)
Lemma 9.1. Let \( \{P_n\} \) be a sequence of probability measures on a Polish space \( X \). Assume that for each \( x \in X \) there exists a sequence of probability measures \( \{\tilde{P}^x_n\} \) weakly convergent to \( \delta_x \) and such that
\[
\lim_{n \to \infty} \frac{1}{n} \text{Ent}(\tilde{P}^x_n | P_n) \leq J(x)
\]
for some \( J: X \to [0, +\infty] \). Then the sequence \( \{P_n\} \) satisfies the large deviation lower bound with rate functional given by \( \text{sc}^{-} J \), the lower semicontinuous envelope of \( J \), i.e.
\[
(\text{sc}^{-} J)(x) := \sup_{U \in \mathcal{X}_x} \inf_{y \in U} J(y)
\]
where $\mathcal{N}_x$ denotes the collection of the open neighborhoods of $x$.

This lemma has been originally proven in [24, Prop. 4.1], see also [33, Prop. 1.2.4].

We first prove the inequalities (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 $\mu_t(x) > \varepsilon$ and $Q_t(y, z) > \varepsilon$ for any $t, x$ 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 $(\text{sc}^- J) = I$.

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

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

Note that the above jump rates are locally integrable and strictly positive for any pair $(y, z) \in E$. We observe that $\mu_t$ satisfies the continuity equation

$$\partial_t \mu_t + \text{div}(\tilde{Q}_t^\mu) = 0.$$  \hfill (9.3)

The symbol $\tilde{Q}_t^\mu$ in (9.2) is defined like in Definition 7.1 by $\tilde{Q}_t^\mu(y, z) := \mu_t(y)\tilde{r}(y, z; t)$. Trivially, $\tilde{Q}_t^\mu = Q$ and therefore (9.3) follows from the definition of $\Lambda_0$. Due to Proposition 7.5 we conclude that $(\mu_t)_{t \geq 0}$ are the marginals 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 $\tilde{P}_n^{(\mu, Q)} := \tilde{\mathbb{P}}_x \circ (\mu^{(n)}, Q^{(n)})^{-1}$. The convergence $\tilde{P}_n^{(\mu, Q)} \to \delta_{(\mu, Q)}$ follows by Lemma 7.3 and the above observation that $\mu_t$ is the marginal of the oscillatory steady state of $\tilde{\mathbb{P}}$.

We first observe that

$$\frac{1}{n} \text{Ent} \left( P_n | P_n \right) \leq \frac{1}{n} \text{Ent} \left( \mathbb{P}_x | [0, nT_0] \right).$$  \hfill (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 entropy.

By a direct computation, using (7.2), we have that the right hand side of (9.4) is given by

$$\tilde{\mathbb{P}}_x \left[ \mu^{(n)}(r - \tilde{r}) + Q^{(n)} \left( \log \frac{\tilde{r}}{r} \right) \right].$$  \hfill (9.5)

Using Lemma 7.3 we get directly that in the limit $n \to +\infty$, (9.5) converges to

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

It remains to prove that $(\text{sc}^- J) = I$. Since $I$ is lower semi-continuous and $I \leq J$ then by definition we have $(\text{sc}^- J) \geq I$. We need to prove the converse inequality. Consider $(\mu, Q) \in \Lambda \cap \Lambda^\circ$. We construct a sequence $(\mu_n, Q_n) \in \Lambda_0$ such that $(\mu_n, Q_n) \to (\mu, Q)$ and moreover $\lim_{n \to +\infty} J(\mu_n, Q_n) \leq I(\mu, Q)$. This implies $(\text{sc}^- J) \leq I$ and allows to conclude the proof.

The sequence $(\mu_n, Q_n)$ 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 $\pi_t(x)$ can be estimated from below by the probability that $\xi_0 = x$ and the Markov chain does not jump in the time interval $[0, t]$. Hence

$$
\pi_t(x) \geq \pi_0(x) \exp\left\{ - \int_0^t r(x; s) ds \right\}.
$$

Due to Assumption (A3) (cf. (2.3)) and since $\pi_0$ is a positive measure, we conclude that $\min_x \inf_{t \in [0, T_0]} \pi_t(x) > 0$. As a byproduct of this bound and again (2.3) we also conclude that $Q_t^I(x, y) = \pi_t(x) r(x, y; t)$ is bounded from below by a positive constant uniformly in $(x, y) \in E$ and $t \in [0, T_0]$. These observations imply that $(\pi, Q^x) \in \Lambda_0$ and therefore that $(\mu_n, Q_n) \in \Lambda_0$.

Since $I$ is convex (cf. Remark 8.3) and $I(\pi, Q^x) = 0$ we have

$$
J(\mu_n, Q_n) = I(\mu_n, Q_n) \leq \left( 1 - \frac{1}{n} \right) I(\mu, Q).
$$

Taking the liminf on both sides of the previous inequality we obtain

$$
\lim_{n \to +\infty} J(\mu_n, Q_n) \leq I(\mu, Q),
$$

and the proof of the lower bound is finished.

**Remark 9.2.** The goodness of the rate functional follows from the exponential tightness in Lemma 8.7 and Lemma 4.1.23 in [14].

10. **Proof of Theorem 3**

Recall the continuous map $\mathcal{J} : \mathcal{M}_a(E \times S_{T_0}) \to \mathcal{M}_a(E \times S_{T_0})$ 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 3.8 it only remains to show that the function

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

indeed equals the r.h.s. of (3.20).

Trivially, if $J = \mathcal{J}(Q)$ with $(\mu, Q) \in \Lambda$, then $(\mu, J) \in \Lambda_a$. Moreover, all elements of $\Lambda_a$ can be obtained in this way. Since $I \equiv +\infty$ on $\Lambda^c$, we conclude that $\tilde{I}(\mu, J) = +\infty$ if $(\mu, J) \notin \Lambda_a$, in agreement with the r.h.s. of (3.20). Hence, from now on we restrict to $(\mu, J) \in \Lambda_a$.

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

$$
J^+(y, z, A) = J^-(z, y, A) \quad \forall A \subset S_{T_0} \text{ measurable}.
$$

Note that by property (v) in Definition 3.8 of $\Lambda_a$, $J^+$ has support included in $E \times S_{T_0}$.

All the flows $Q \in \mathcal{M}_+(E \times S_{T_0})$ such that $\mathcal{J}(Q) = J$ can be characterized by the decomposition $Q = J^+ + S$, where $S$ is an arbitrary element of $\mathcal{M}_+(E \times S_{T_0})$ 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 $S = S(\mu)$ the space of measures $S \in \mathcal{M}_+(E \times S_{T_0})$ such that $S = S_{dt}$, $S_t \in \mathbb{R}_+^E$,
\[
\begin{cases}
S_t(y, z) = S_t(z, y) & \text{if } (y, z) \in E \text{ and } (z, y) \in E,
S_t(y, z) = 0 & \text{if } (y, z) \in E \text{ and } (z, y) \notin E.
\end{cases}
\]
and, given $(y, z) \in E$, if $\mu_t(y) = 0$ then $S_t(y, z) = 0$ for a.e. $t$.

Recall that we restrict to $(\mu, J) \in \Lambda$. 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 $S \in S$.

Due to the previous observations we have
\[
\hat{I}(\mu, J) = \inf_{S \in S} I(\mu, J^+ + S) = \inf_{S \in S} \sum_{(y, z) \in E} \int_0^{T_0} \Phi(J_t^+(y, z) + S_t(y, z), \mu_t(y) r_t(y, z)) dt,
\]
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 addenda associated to $(y, z)$ and to $(z, y)$ (if $(z, y) \in E$, otherwise one restricts only to the addendum associated to $(y, z)$).

To this aim, given $(v, w) \in E$ we set
\[
Q_t^J(\mu, v, w) := \frac{J_t(v, w) + \sqrt{J_t^2(v, w) + 4\mu_t(v)\mu_t(w)r_t(v, w)r_t(w, v)}}{2}.
\]

Case 1. For $(y, z) \in E$ with $(z, y) \notin E$ we know that $S_t(y, z) = 0$, $Q_t(y, z) = J_t^+(y, z) = J_t(y, z)$. Therefore, for all $t \in [0, T_0]$, we have
\[
\Phi(J_t^+(y, z) + S_t(y, z), \mu_t(y) r_t(y, z)) = \Phi(Q_t^J(\mu, y, z), \mu_t(y) r_t(y, z)).
\]

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
\[
\Phi(J_t^+(y, z) + S_t(y, z), \mu_t(y) r_t(y, z)) + \Phi(J_t^+(z, y) + S_t(z, y), \mu_t(z) r_t(z, y)),
\]
when varying the parameter $S_t(y, z) = S_t(z, y)$ in $\mathbb{R}_+$. We define
\[
s_t := S_t(y, z) = S_t(z, y), \quad j_t^+ := J_t^+(y, z), \quad j_t^- := J_t^-(y, z) = -J_t^+(z, y).
\]

Subcase 2.a. Supposing $\mu_t(y) > 0$ and $\mu_t(z) > 0$, by definition of $\Phi$ we have to minimise (cf. \[(10.3)\])
\[
\inf_{s_t \in \mathbb{R}_+} \left\{ \left( j_t^+ + s_t \right) \log \frac{j_t^+ + s_t}{\mu_t(y) r_t(y, z)} + \left( j_t^- + s_t \right) \log \frac{j_t^- + s_t}{\mu_t(z) r_t(z, y)} + \mu_t(y) r_t(y, z) + \mu_t(z) r_t(z, y) - j_t^+ - j_t^- - 2s_t \right\}.
\]
By simple computations one gets that the minimizer is given by
\[
s_t = \frac{-\left( j_t^+ + j_t^- \right) + \sqrt{(j_t^+ - j_t^-)^2 + 4\mu_t(y)\mu_t(z)r_t(y, z)r_t(z, y)}}{2}.
\]
We point out that \( s_t > 0 \) since \( \min(j_t^+, j_t^-) = 0 \). It then follows that the infimum in (10.3) equals
\[
\Phi\left(Q_t^{j_\mu}(y, z), \mu_t(y)r_t(y, z)\right) + \Phi\left(Q_t^{j_\mu}(z, y), \mu_t(z)r_t(z, y)\right). \tag{10.5}
\]

**Subcase 2.b.** If \( \mu_t(y) = 0 \) and \( \mu_t(z) > 0 \), then by Property (iv) in Definition 10.3 and by Definition 10.1 of \( S \) for a.e. \( t \) we have \( j_t^+ = 0 = s_t \). In this case, for a.e. \( t \) the contribution (10.4) equals
\[
\sum_{j = -1}^0 j_t^- \log \frac{j_t^-}{\mu_t(z)r_t(z, y)} + \mu_t(z)r_t(z, y) - j_t^- \tag{10.6}
\]
which again equals (10.5).

**Subcase 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
\[
\hat{I}(\mu, J) = \int_0^{T_0} I_t(\mu_t, Q_t^{j_\mu}) \, dt. \tag{10.7}
\]
Finally, the derivation of (3.22) from the above formula can be done as in [6] (cf. Theorem 6.1 there) by correcting the conclusion there. Let us give more comments. As for [6] Eq. (6.6) we have
\[
\Psi(J_t(y, z), J_t^\mu(y, z); a_t^\mu(y, z)) = \Phi\left(Q_t^{j_\mu}(y, z), \mu_t(y)r_t(y, z); t\right) + \Phi\left(Q_t^{j_\mu}(z, y), \mu_t(z)r_t(z, y); t\right)
\]
if both \((y, z)\) and \((z, y)\) belong to \( E \). Hence in this case we have
\[
\frac{1}{2}\left\{\Psi(J_t(y, z), J_t^\mu(y, z); a_t^\mu(y, z)) + \Psi(J_t(z, y), J_t^\mu(z, y); a_t^\mu(z, y))\right\} = \\
\Phi\left(Q_t^{j_\mu}(y, z), \mu_t(y)r_t(y, z); t\right) + \Phi\left(Q_t^{j_\mu}(z, y), \mu_t(z)r_t(z, y); t\right). \tag{10.8}
\]
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
\[
\Psi(J_t(y, z), J_t^\mu(y, z); a_t^\mu(y, z)) = \Phi\left(Q_t^{j_\mu}(y, z), \mu_t(y)r_t(y, z); t\right). 
\]
On the other hand, we have \( a_t^\mu(z, y) = 0 \), \( J_t(z, y) = -J_t^\mu(y, z) \leq 0 \) and \( J_t^\mu(z, y) = -J_t^\mu(y, z) \leq 0 \). Hence, by definition of \( \Psi \), we have
\[
\Psi(J_t(z, y), J_t^\mu(z, y); a_t^\mu(z, y)) = \Psi(J_t(y, z), J_t^\mu(y, z); a_t^\mu(y, z)).
\]
Since moreover \( \Phi(Q_t^{j_\mu}(z, y), \mu_t(z)r_t(z, y); t) = \Phi(0, 0) = 0 \), also in this case we have (10.8). By symmetry we conclude that (10.8) holds for any \((y, z) \in E \). As a byproduct of the above observation, (3.22) and (10.7), we get (3.22).

We conclude by discussing goodness and convexity of \( \hat{I} \). Goodness follows from the goodness of \( I \) by application of the contraction principle. On the other hand, by (3.18), \( \hat{I}(\mu, Q) \) equals the infimum of the convex rate functional \( I \) on a suitable affine subspace, thus implying that \( \hat{I} \) 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 $(\mu, Q)$ and $((\theta \mu, \theta Q)$ belong to $\Lambda$ we can write

\[ I(\theta \mu, \theta Q; r) - I(\mu, Q; r) = \sum_{y,z} \int_{0}^{T_0} ds \left[ -Q_s(y, z) \log \frac{Q_s(y, z)}{\mu_s(y)r(y, z; 0)} + Q_{T_0-s}(z, y) \log \frac{Q_{T_0-s}(z, y)}{\mu_{T_0-s}(y)r(y, z; 0)} \right] \]

\[ + \sum_y \int_{0}^{T_0} ds \left[ -\sum_z Q_s(y, z) - \mu_s(y)r(y, s) \right] \]

\[ = \sum_{y,z} \int_{0}^{T_0} ds \left[ -Q_s(y, z) \log \frac{Q_s(y, z)}{\mu_s(y)r(y, z; 0)} + Q_s(z, y) \log \frac{Q_s(z, y)}{\mu_s(y)r(y, z; 0)} \right] \]

\[ + \sum_y \int_{0}^{T_0} ds \left[ -\sum_z Q_s(y, z) - \mu_s(y)r(y, s) \right] \].

Therefore we have

\[ I(\theta \mu, \theta Q; r) - I(\mu, Q; r) = \int_{0}^{T_0} ds \left[ \sum_{y,z} Q_s(y, z) \log \frac{\mu_s(y)r(y, z; s)}{\mu_s(z)r(y, z; T_0 - s)} + \sum_y \mu_s(y) (-r(y, s) + r(y, T_0 - s)) \right]. \]

(11.1)

On the other hand we have

\[ \int_{0}^{T_0} ds \sum_{y,z} Q_s(y, z) \log \frac{\mu_s(y)}{\mu_s(z)} = \int_{0}^{T_0} ds \sum_y \log(\mu_s(x)) \sum_z (Q_s(y, z) - Q_s(z, y)). \]

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

\[ \int_{0}^{T_0} ds \sum_{y,z} Q_s(y, z) \log \frac{\mu_s(y)}{\mu_s(z)} = \sum_y \int_{0}^{T_0} ds \log(\mu_s(y)) \partial_s \mu_s(y) = 0. \]

(11.2)

As a byproduct of (11.1) and (11.2) one gets (5.2).

11.2. Proof of (5.3). Since $(\mu, Q) \in \Lambda$ we can write

\[ I(\theta \mu, \theta Q; r^R) - I(\mu, Q; r) = \sum_{y,z} \int_{0}^{T_0} ds \left[ -Q_s(y, z) \log \frac{Q_s(y, z)}{\mu_s(y)r(y, z; 0)} + Q_{T_0-s}(z, y) \log \frac{Q_{T_0-s}(z, y)}{\mu_{T_0-s}(y)r(y, z; 0)} \right] \]

\[ + \sum_y \int_{0}^{T_0} ds \left[ -\sum_z Q_s(y, z) - \mu_s(y)r(y, s) \right] \].

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

\[ I(\theta \mu, \theta Q; r^R) - I(\mu, Q; r) = \sum_{y,z} \int_{0}^{T_0} ds Q_s(y, z) \log \frac{\mu_s(y)r(y, z; s)}{\mu_s(z)r(y, z; s)} \].

By (11.2) we can write the above r.h.s. as $\sum_{y,z} \int_{0}^{T_0} ds Q_s(y, z) \log \frac{r(y, z; s)}{r(y, z; 0)} = S_{tot}(\mu; r)$.
11.3. Proof of (5.4). Since \((\mu, Q) \in \Lambda\) we can write

\(I(\theta \mu, \theta Q; r^{\text{DR}}) - I(\mu, Q; r)\)

\[
= \sum_{y, z} \int_{T_0} ds \left[ -Q_s(y, z) \log \frac{Q_s(y, z)}{\mu_s(y) w_s^{-1}(y)} \right] + Q_{T_0 - s}(z, y) \log \frac{Q_{T_0 - s}(z, y)}{\mu_s(y) w_s^{-1}(y) r(y; T_0 - s) w_{T_0 - s}(z)}
\]

\[
+ \sum_y \int_{T_0} ds \left[ -\sum_z Q_s(y, z) - \mu_s(y) r(y; s) \right]
\]

\[
= \sum_{y, z} \int_{T_0} ds \left[ -Q_s(y, z) \log \frac{Q_s(y, z)}{\mu_s(y) w_s^{-1}(y)} \right] + Q_{T_0 - s}(z, y) \log \frac{Q_{T_0 - s}(z, y)}{\mu_s(y) w_s^{-1}(y) r(y; T_0 - s)}
\]

\[
+ \sum_y \int_{T_0} ds \left[ -\sum_z Q_s(y, z) - \mu_s(y) r(y; s) \right]
\]

\[
= \int_0^{T_0} ds \sum_{y, z} Q_s(y, z) \log \frac{\mu_s(y) w_s^{-1}(y)}{\mu_s(z) w_s^{-1}(z)}
\]

\[
= \int_0^{T_0} ds \sum_{y, z} Q_s(y, z) \log \frac{w_s(z)}{w_s(y)}.
\]

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

\[
\int_0^{T_0} ds \sum_{y, z} Q_s(y, z) \log \frac{w_s(z)}{w_s(y)} = \int_0^{T_0} ds \sum_z \log(w_s(z)) \sum_y (Q_s(y, z) - Q_s(z, y))
\]

\[
= \sum_z \int_0^{T_0} ds \log(w_s(z)) \partial_s \mu_s(z) = - \sum_z \int_0^{T_0} ds \mu_s(z) \partial_s \log(w_s(z)) = S_{\text{ex}}(\mu; r).
\]

11.4. Proof of (5.6), (5.7) and (5.7). These last three identities follows 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 of \(\Lambda\) and to use the identities \(S_{\text{naive}}(\theta \mu, \theta Q; r) = -S_{\text{naive}}(\mu, Q; r), S_{\text{tot}}(\theta Q; r^R) = -S_{\text{tot}}(Q; r), S_{\text{ex}}(\theta \mu; r^{\text{DR}}) = -S_{\text{ex}}(\mu; r)\). For the last identity we observe that the accompanying measure \(w_s^{\text{DR}}\) associated to the rates \(r^{\text{DR}}(\cdot, \cdot; s)\) equals \(w_{T_0 - s}\).

Acknowledgements. A. Faggionato and D. Gabrielli thank the Institut 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., Seifert U.; Thermodynamic Uncertainty Relation for Biomolecular Processes. Phys. Rev. Lett. 114, 158101 (2015).

[2] Billingsley P.; Convergence of Probability Measures. Second ed., in: Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1999.
[3] Bertini L., De Sole A., Gabrielli D., Jona-Lasinio G., Landim C.; *Macroscopic fluctuation theory*. Rev. Modern 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 Processes 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. Statist. 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*. Nature Physics 8, 143–146 (2012).
[8] Chernyak V., Chertkov M., Jarzynski C.; *Path-integral analysis of fluctuation theorems for general Langevin processes*. J. Stat. Mech. P08001, (2006).
[9] 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).
[10] Chetrite R., Gawedzki K.; *Fluctuation relations for diffusion processes*. Commun. Math. Phys. 282, 469–518 (2008).
[11] Crooks G.E.; *Non-equilibrium measurements of free energy differences for microscopically reversible Markovian systems*. J. Stat. Phys. 90, 1481–1487 (1998).
[12] Davis M. H. A.; *Markov models and optimization*. London, Chapman & Hall, 1993.
[13] Dembo A., Zeitouni O.; *Large Deviation Techniques and Applications*. Second ed. Springer, New York, 1998.
[14] 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).
[15] De la Fortelle A.; *Large Deviation Principle for Markov Chains in Continuous Time*. Prob. Inf. Transm. 37, 120 (2001).
[16] Faggionato A., Gabrielli D., Ribezzi Crivellari M.; *Non-equilibrium thermodynamics of piecewise deterministic Markov processes*. J. Stat. Phys. 137 259–304 (2009).
[17] Gingrich T., Horowitz J., Perunov N., England J.; *Dissipation bounds all steady-state current fluctuations*. Phys Rev Lett 116 (12) 120601 (2016).
[18] Gingrich T., Rotskoff G., Horowitz J.; *Inferring dissipation from current fluctuations*. J Phys A 50 (18) 184004 (2017).
[19] Ge H., Jiang D.-Q., Qian M.; *Reversibility and entropy production of inhomogeneous Markov chains*. Journal of Applied Probability 43, 1028–1043 (2006).
[20] Harris R.J., Schütz J.M.; *Fluctuation theorems for stochastic dynamics*. J. Stat. Mech. P07020 (2007).
[21] Hanggi P., Thomas H.; *Stochastic processes: time evolution, symmetries and linear response*. Phys. Rep. 88, 207–319 (1982).
[22] Hatano, T., Sasa, S.; *Steady-state thermodynamics of Langevin systems*. Phys. Rev. Lett. 86, 3463–3466 (2001).
[23] Höpfner R., Kutoyants Y.; *Estimating discontinuous periodic signals in a time inhomogeneous diffusion*. Statist. Inference Stoch. Proc. 13, 193–230 (2010).
[24] Jensen L.H.; *Large deviations of the asymmetric simple exclusion process in one dimension*. Ph.D. Thesis, Courant Institute NYU (2000).
[25] Kaiser M., Jack R.L., Zimmer J.; *Canonical structure and orthogonality of forces and currents in irreversible Markov chains*. Preprint, arXiv:1705.01453.
[26] Kesidis G., Walrand J.; *Relative entropy between Markov transition rate matrices*, IEEE Trans. Info. Theo. 39, 10561057 (1993).
[27] Kusuoka S., Kuwada K., Tamura Y.; *Large deviation for stochastic line integrals as Lp–currents*. Probab. Theory Relat. Fields 147, 649 (2010).
[28] Kipnis C., Landim C.; *Scaling limits of interacting particle systems*. Springer-Verlag, Berlin, 1999.
[29] Lazarescu A.; *The physicist’s companion to current fluctuations: one-dimensional bulk-driven lattice gases*. J. Phys. A 48, 503001 (2015).
[31] 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).
[32] Maes C.; The fluctuation theorem as a Gibbs property. J. Stat. Phys. 95, 367–392 (1999).
[33] Maes C., Netocny K., Wynants B.; Steady state statistics of driven diffusions. Physica A 387, 2675 (2008).
[34] Maes C., Netocny K.; The canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states. Europhys. Lett. 82, 30003 (2008).
[35] Mariani M.; A Γ–convergence approach to large deviations. To appear in Ann. Sc. Norm. Super. Pisa Cl. Sci.
[36] Martinez I.A., Roldán É, Dinis L., Parrondo J.M.R., Rica R.A.; Brownian Carnot engine. Nature Physics 12, 67-70 (2016).
[37] McNamara B.; Wiesenfeld K.; Theory of stochastic resonance. Phys. Rev. A 39, 4854 (1989).
[38] Mörters P., Peres Y.; Brownian motion. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge 2010.
[39] Norris J.R.; Markov chains. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge 1999.
[40] Pietzonka P., Barato, A., Seifert U.; Universal bounds on current fluctuations. Physical Review E 93 (5), 052145 (2016).
[41] Pietzonka P., Barato A., Seifert U.; Affinity-and topology-dependent bound on current fluctuations. Journal of Physics A: Mathematical and Theoretical 49 (34), 34LT01 (2016).
[42] Pietzonka P., Barato A.; Stochastic thermodynamics of periodically driven systems: fluctuation theorem for currents and unification of two classes. Preprint, arXiv 1707.08100
[43] Proesmans K.; Cleuren B.; Van den Broeck C.; Linear stochastic thermodynamics for periodically driven systems. J. Stat. Mech. 023202 (2016).
[44] Ray S., Barato A.; Stochastic thermodynamics of periodically driven systems: fluctuation theorem for currents and unification of two classes. Preprint, arXiv 1707.08100
[45] 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).
[46] Singh N.; Onsager-Machlup theory and work fluctuation theorem for a harmonically driven Brownian particle. J. Stat. Phys. 131, 405–414 (2008).
[47] Singh N., Wynants B.; Dynamical fluctuations for periodically driven diffusions. J. Stat. Mech. P03007 (2010).
[48] Schüller 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).
[49] Sinitsyn N.A., Akimov A., Chernyak V.Y.; Supersymmetry and fluctuation relations for currents in closed networks Phys. Rev. E 83, 021107 (2011).
[50] Verley G., Van den Broeck C., Esposito M.; Modulated two-level system: Exact work statistics. Phys. Rev. E 88, 032137 (2013).
