Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 125 (2015) 2786–2819

stochastic processes and their applications

www.elsevier.com/locate/spa

Flows, currents, and cycles for Markov chains: Large deviation asymptotics

Lorenzo Bertini [a], Alessandra Faggionato [a,*], Davide Gabrielli [b]

[a] *Dipartimento di Matematica, Università di Roma 'La Sapienza', P.le Aldo Moro 2, 00185 Roma, Italy*
[b] *DISIM, Università dell'Aquila, Via Vetoio, 67100 Coppito, L'Aquila, Italy*

Received 25 July 2014; received in revised form 3 February 2015; accepted 3 February 2015
Available online 24 February 2015

Abstract

We consider a continuous time Markov chain on a countable state space. We prove a joint large deviation principle (LDP) of the empirical measure and current in the limit of large time interval. The proof is based on results on the joint large deviations of the empirical measure and flow obtained in Bertini et al. (in press). By improving such results we also show, under additional assumptions, that the LDP holds with the strong L^1 topology on the space of currents. We deduce a general version of the Gallavotti–Cohen (GC) symmetry for the current field and show that it implies the so-called fluctuation theorem for the GC functional. We also analyze the large deviation properties of generalized empirical currents associated to a fundamental basis in the cycle space, which, as we show, are given by the first class homological coefficients in the graph underlying the Markov chain. Finally, we discuss in detail some examples.

MSC: primary 60F10; 60J27; secondary 82C05

Keywords: Markov chain; Large deviations; Empirical flow; Empirical current; Cellular homology; Gallavotti–Cohen fluctuation theorem

* Corresponding author.
E-mail addresses: bertini@mat.uniroma1.it (L. Bertini), faggiona@mat.uniroma1.it (A. Faggionato), gabriell@univaq.it (D. Gabrielli).

http://dx.doi.org/10.1016/j.spa.2015.02.001

1. Introduction

We consider a continuous time Markov chain on a countable (finite or infinite) state space V with transition rates $r(\cdot, \cdot)$. We assume that the chain is ergodic and positive recurrent, so that it admits a unique invariant probability distribution π.

A natural observable is given by the *empirical measure* μ_T, which accounts for the fraction of time spent on the various states up to time T. As $T \to \infty$, μ_T converges to π. The large deviation principle for the family $\{\mu_T\}$ is the classical Donsker–Varadhan theorem [13]. Other natural observables are the *empirical flow* Q_T and *empirical current* J_T, which respectively account for the total numbers of jumps and for the net flow between pairs of states per unit of time. In particular, given two states $y, z \in V$, it holds $J_T(y, z) = Q_T(y, z) - Q_T(z, y)$. As $T \to \infty$, $Q_T(y, z)$ and $J_T(y, z)$ respectively converge to $\pi(y)r(y, z)$ and $\pi(y)r(y, z) - \pi(z)r(z, y)$. The large deviation principle for the family $\{(\mu_T, Q_T)\}$ is proven in [5].

The interest for these observables comes from several applications. We mention some of them, mainly related to the concept of work, to the Gallavotti–Cohen functional and to the concept of activity in kinetically constrained spin systems.

When the Markov chain models the stochastic dynamics of a physical particle in presence of an external field and thermal noise, the work done by the field can be expressed in terms of the empirical current. When modeling biochemical systems, the state describes both the mechanical and the chemical configuration. One is then interested in the work done both by the applied mechanical force and the chemical one, in which the latter is induced by differences in the chemical potentials. In both cases the work is a linear function of the empirical current. Significant examples are biochemical systems given by single molecules like molecular motors [27].

In out-of-equilibrium statistical mechanics a much studied observable is the Gallavotti–Cohen functional W_T. It is defined as follows [20]: e^{-TW_T} is the Radon–Nikodym derivative of the time-reversed stationary process $\mathbb{P}^*_\pi$ w.r.t. the stationary process itself $\mathbb{P}_\pi$ in the time window $[0, T]$. It follows that W_T accounts for the irreversibility of the stochastic dynamics and its expectation w.r.t. $\mathbb{P}_\pi$ is the relative entropy of $\mathbb{P}_\pi$ w.r.t. $\mathbb{P}^*_\pi$ per unit of time. By a straightforward computation, it turns out that W_T is a linear function of the empirical current apart boundary terms. When the state space is finite, the large deviation principle for $\{W_T\}$ has been derived in [20] by the Gärtner–Ellis theorem. The so-called *fluctuation theorem* (or *Gallavotti–Cohen symmetry*) is then the identity $\iota(u) - \iota(-u) = -u$ satisfied by the corresponding rate function $\iota : \mathbb{R} \to \mathbb{R}_+$.

For kinetically constrained spin systems, see [7] and references therein, the empirical flow is a relevant observable and its large deviation properties exhibit peculiar and rich features. More precisely, given a system of N spins, the N-normalized total number of jumps per unit time (also called *activity*) has a nontrivial second order LDP in the limit $T \to \infty$ and afterwards $N \to \infty$. We point out that the above activity is proportional to the total mass of the empirical flow.

Starting from the results in [5], in this paper we derive the large deviation principle for the family $\{(\mu_T, J_T)\}$ (Theorem 6.1). By contraction, we then deduce the large deviation principle for the Gallavotti–Cohen functional and show the rate function ι satisfies the Gallavotti–Cohen symmetry (Theorem 8.1). We remark that this derivation yields an explicit variational representation of ι, while the derivation via Gärtner–Ellis theorem gives a spectral characterization [20]. For infinite state spaces there are however some technical issues that are best exemplified in the case of a single particle performing a random walk on $\mathbb{Z}^d$ with confining potential U and external field F. Since the result in [5] is proven by using the bounded weak* topology for the empirical flow, the contraction can be performed only when the external field vanishes at infinity. On the other hand, a natural condition is that F is bounded. To overcome the requirement of F vanishing at infinity, we prove the large deviation principle for $\{(\mu_T, Q_T)\}$ in the strong L^1 topology for the empirical flow (Theorem 5.2) under (needed) additional conditions in the general setting. As further reinforcement of the results of [5] we also show that some technical assumption there can be dropped (see Proposition 4.1).

We continue our investigation of Gallavotti–Cohen type symmetries. Consider the transition graph G, with vertex set V, of the Markov chain. For biochemical models, as explained in Section 9, the work of the mechanical/chemical forces can be expressed in terms of the homological coefficients of the trajectory in a suitable basis of the first cellular homology class $H_1(G;\mathbb{R})$ of G [26]. For finite state space, the analysis of the large deviations of the homological coefficients and the related Gallavotti–Cohen symmetry has been deduced in [1,14] via Gärtner–Ellis theorem. We extend this result to infinite state space emphasizing the relationship of the homological coefficients with the empirical current (Theorem 9.4). To perform this extension in Section 9 we present a self-contained overview on concepts from graph theory and cycle spaces with emphasis on the infinite case setting. We finally point out that the Gallavotti–Cohen symmetry both for the Gallavotti–Cohen functional and the homological coefficients is a consequence of a general symmetry of the rate functional for $\{(\mu_T, J_T)\}$ (Theorem 7.1).

Finally, in Section 10 we discuss several examples in which some rate functionals can be computed explicitly.

We conclude with further bibliographical remarks. In the context of finite state space, the joint LDPs for $\{(\mu_T, Q_T)\}$ and $\{(\mu_T, J_T)\}$ have been discussed in [16,21,23]. See also [3] for a perturbative expansion in the context of non-equilibrium statistical mechanics. For countable state spaces, a weak form of joint LDP for $\{(\mu_T, Q_T)\}$ is derived in [10]. In [17] large deviation properties of periodic random walks on crystal lattices are analyzed and related to the first homology group of the finite quotient graph. The joint LDP for $\{(\mu_T, J_T)\}$ of a Brownian motion on a compact Riemannian manifold is proved in [18,19]. See also the discussion in [22] for diffusions on the torus $\mathbb{T}^d$ and on $\mathbb{R}^d$ with a confining potential.

2. Basic setting

We consider a continuous time Markov chain ξ_t, $t \in \mathbb{R}_+$ on a countable (finite or infinite) state space V. The Markov chain is defined in terms of the *jump rates* $r(x, y)$, $x \neq y$ in V, from which one derives the holding times and the jump chain [25].

The basic assumptions on the chain are the following:

(A1) for each $x \in V$, $r(x) := \sum_{y\in V} r(x, y)$ is finite;
(A2) for each $x \in V$ the Markov chain ξ_t^x starting from x has no explosion a.s.;
(A3) the Markov chain is irreducible, i.e. for each $x, y \in V$ and $t > 0$ the event $\{\xi_t^x = y\}$ has strictly positive probability;
(A4) there exists a unique invariant probability measure, that is denoted by π.

By assumption (A1) the holding time at $x \in V$ is a well defined exponential random variable of parameter $r(x)$. As in [25], by invariant probability measure π we mean a probability measure on V such that

$$\sum_{y\in V} \pi(x)\, r(x, y) = \sum_{y\in V} \pi(y)\, r(y, x) \quad \forall x \in V \tag{2.1}$$

where we understand $r(x, x) = 0$. We refer to Section 4 for a discussion on the above assumptions (A1), …,(A4) and their relation with Condition $C(\sigma)$ introduced in the next section. We only recall that $\pi(x) > 0$ for all $x \in V$ and that the Markov chain starting with distribution π is stationary (i.e. is left invariant by time-translations).

We consider V endowed with the discrete topology and the associated Borel σ-algebra given by the collection of all the subsets of V. Given $x \in V$, the distribution of the Markov chain ξ_t^x starting from x, is a probability measure on the Skorohod space of càdlàg paths $D(\mathbb{R}_+; V)$ that we denote by $\mathbb{P}_x$. The expectation with respect to $\mathbb{P}_x$ is denoted by $\mathbb{E}_x$. In the sequel we consider $D(\mathbb{R}_+; V)$ equipped with the canonical filtration, the canonical coordinate in $D(\mathbb{R}_+; V)$ is denoted by X_t. The set of probability measures on V is denoted by $\mathcal{P}(V)$ and it is considered endowed with the topology of weak convergence and the associated Borel σ-algebra.

In view of Assumptions (A1)–(A4) the ergodic theorem holds: for any bounded function $f : V \to \mathbb{R}$ and for any $x \in V$

$$\lim_{T\to+\infty} \frac{1}{T} \int_0^T dt\, f(\xi_t) = \langle \pi, f \rangle \quad \mathbb{P}_x\text{-a.s.} \tag{2.2}$$

where $\langle \pi, f \rangle$ denotes the expectation of f with respect to π.

2.1. *Empirical measure and empirical flow*

Given $T > 0$ the *empirical measure* $\mu_T : D(\mathbb{R}_+; V) \to \mathcal{P}(V)$ is defined by

$$\mu_T\,(X) = \frac{1}{T} \int_0^T dt\, \delta_{X_t},$$

where δ_y denotes the point mass at y. By the ergodic theorem the sequence of probabilities $\{\mathbb{P}_x \circ \mu_T^{-1}\}_{T>0}$ on $\mathcal{P}(V)$ converges to δ_π.

We denote by E the (countable) set of ordered edges in V with strictly positive jump rate, i.e.

$$E := \{(y, z) \in V \times V : r(y, z) > 0\},$$

by $L^1(E)$ the collection of absolutely summable functions on E and by $\|\cdot\|$ the associated L^1-norm. The set of positive elements in $L^1(E)$ is denoted by $L^1_+(E)$. Note that, since V has the discrete topology and is countable, any path in $D(\mathbb{R}_+; V)$ has a locally finite number of jumps. In particular, for each $T > 0$ we can define the *empirical flow* as the map $Q_T : D(\mathbb{R}_+; V) \to L^1_+(E)$ given by

$$Q_T(y, z)\,(X) := \frac{1}{T} \sum_{0\le t\le T} \mathbb{1}\,(X_{t^-} = y,\ X_t = z) \quad (y, z) \in E, \tag{2.3}$$

where, in general, $\mathbb{1}(A)$ denotes the characteristic function of A. Namely, $T Q_T(y, z)$ gives the number of jumps from y to z in the time interval $[0, T]$.

Elements of $L^1_+(E)$ will be denoted by Q and called *flows*. Given a flow Q we let its *divergence* $\mathrm{div}Q : V \to \mathbb{R}$ be the pointwise difference between the outgoing flow and the ingoing one, namely

$$\mathrm{div}Q\,(y) = \sum_{z:\,(y,z)\in E} Q(y, z) - \sum_{z:\,(z,y)\in E} Q(z, y), \quad y \in V. \tag{2.4}$$

Observe that the divergence maps $L^1_+(E)$ to $L^1(V)$. To each probability $\mu \in \mathcal{P}(V)$ such that $\langle \mu, r \rangle < +\infty$ we associate the flow Q^μ defined by

$$Q^\mu(y, z) := \mu(y)\, r(y, z) \quad (y, z) \in E. \tag{2.5}$$

Note that Q^μ has vanishing divergence if and only if μ is invariant, i.e. $\mu = \pi$.

By the ergodic theorem and a martingale argument (cf. [5]) one can show that for each $x \in V$ and $(y, z) \in E$ the sequence of real random variables $Q_T(y, z)$ converges as $T \to +\infty$ to $Q^\pi(y, z)$ in probability with respect to $\mathbb{P}_x$.

3. Joint large deviations for the empirical measure and flow

In this section we recall the main results of [5]. The space $L^1_+(E)$ is endowed with the bounded weak* topology [24], which is defined as follows. Let $C_0(E)$ be the space of functions $f : E \to \mathbb{R}$ vanishing at infinity, endowed with the uniform norm. Then its dual space is given by $L^1(E)$ endowed with the strong topology (i.e. the topology determined by the L^1-norm). A basis of the bounded weak* topology on $L^1(E)$ is then given by the sets

$$\{q \in L^1(E) : \langle q - \bar{q}, f_n\rangle < 1 \ \forall n \ge 1\}$$

as $\bar{q}$ varies among $L^1(E)$ and $(f_n)_{n\ge1}$ varies among the sequences in $C_0(E)$ converging to 0 in uniform norm. In general, given $q \in L^1(E)$ and $f \in C_0(E)$, we set $\langle q, f\rangle := \sum_{e\in E} q(e) f(e)$. Finally, the bounded weak* topology on $L^1_+(E)$ is the inherited subspace topology on $L^1_+(E) \subset L^1(E)$, when $L^1(E)$ itself is endowed with the above defined bounded weak* topology.

One can prove (cf. [24][Corollary 2.7.4]) that a subset $W \subset L^1(E)$ is open in the bounded weak* topology if and only if for each $\ell > 0$ the set $\{q \in W : \|q\|_1 \le \ell\}$ is open in the ball $\{q \in L^1(E) : \|q\|_1 \le \ell\}$ endowed with the weak* topology inherited from $L^1(E)$. We recall that the weak* topology of $L^1(E)$ is the weakest topology such that the map $L^1(E) \ni q \to \langle q, f\rangle \in \mathbb{R}$ is continuous for any map $f \in C_0(E)$. When E is finite, the bounded weak* topology coincides with the strong topology. If E is infinite then the former is weaker than the latter and cannot be metrized.

We can now recall the LDP proved in [5]. We start from the assumptions. To this aim, given $f : V \to \mathbb{R}$ such that $\sum_{y\in V} r(x, y)\,|f(y)| < +\infty$ for each $x \in V$, we denote by $Lf : V \to \mathbb{R}$ the function defined by

$$Lf(x) := \sum_{y\in V} r(x, y)\big[f(y) - f(x)\big], \quad x \in V. \tag{3.1}$$

Condition $C(\sigma)$. Given $\sigma \in \mathbb{R}_+$ we say that Condition $C(\sigma)$ holds if there exists a sequence of functions $u_n : V \to (0, +\infty)$ satisfying the following requirements:

(i) For each $x \in V$ and $n \in \mathbb{N}$ it holds $\sum_{y\in V} r(x, y) u_n(y) < +\infty$.

(ii) The sequence u_n is uniformly bounded from below. Namely, there exists $c > 0$ such that $u_n(x) \ge c$ for any $x \in V$ and $n \in \mathbb{N}$.

(iii) The sequence u_n is uniformly bounded from above on compacts. Namely, for each $x \in V$ there exists a constant C_x such that for any $n \in \mathbb{N}$ it holds $u_n(x) \le C_x$.

(iv) Set $v_n := -Lu_n/u_n$. The sequence $v_n : V \to \mathbb{R}$ converges pointwise to some $v : V \to \mathbb{R}$.

(v) The function v has compact level sets. Namely, for each $\ell \in \mathbb{R}$ the level set $\{x \in V : v(x) \le \ell\}$ is finite.

(vi) There exists a positive constant C such that $v \ge \sigma\, r - C$.

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

$$\Phi(q, p) := \begin{cases} q \log \dfrac{q}{p} - (q - p) & \text{if } q, p \in (0, +\infty) \\ p & \text{if } q = 0,\ p \in [0, +\infty) \\ +\infty & \text{if } p = 0 \text{ and } q \in (0, +\infty). \end{cases} \tag{3.2}$$

For $p > 0$, $\Phi(\cdot, p)$ is a nonnegative strictly convex function and is zero only at $q = p$. Indeed, since $\Phi(q, p) = \sup_{s\in\mathbb{R}} \{qs - p(e^s - 1)\}$, Φ is the rate function 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.

Finally, we let $I : \mathcal{P}(V) \times L^1_+(E) \to [0, +\infty]$ be the functional defined by

$$I(\mu, Q) := \begin{cases} \sum_{(y,z)\in E} \Phi\big(Q(y, z), Q^\mu(y, z)\big) & \text{if } \operatorname{div} Q = 0,\ \langle \mu, r\rangle < +\infty \\ +\infty & \text{otherwise.} \end{cases} \tag{3.3}$$

Remark 3.1. As stated in [5, Remark 2.6] the above condition $\langle \mu, r\rangle < +\infty$ can be removed, since the series in (3.3) diverges if $\langle \mu, r\rangle = +\infty$.

Theorem 3.2 *([5, Theorem 2.7]). Endow $\mathcal{P}(V)$ with the weak topology and $L^1_+(E)$ with the bounded weak* topology. Assume the Markov chain satisfies* (A1)–(A4) *and ConditionC(σ) with $\sigma > 0$. Then, as $T \to +\infty$, the sequence of probability measures $\{\mathbb{P}_x \circ (\mu_T, Q_T)^{-1}\}$ on $\mathcal{P}(V) \times L^1_+(E)$ satisfies a LDP with good and convex rate function I. Namely, for each closed set $\mathcal{C} \subset \mathcal{P}(V) \times L^1_+(E)$, and each open set $\mathcal{A} \subset \mathcal{P}(V) \times L^1_+(E)$, it holds for each $x \in V$*

$$\varlimsup_{T\to+\infty} \frac{1}{T} \log \mathbb{P}_x\Big((\mu_T, Q_T) \in \mathcal{C}\Big) \le - \inf_{(\mu,Q)\in\mathcal{C}} I(\mu, Q), \tag{3.4}$$

$$\varliminf_{T\to+\infty} \frac{1}{T} \log \mathbb{P}_x\Big((\mu_T, Q_T) \in \mathcal{A}\Big) \ge - \inf_{(\mu,Q)\in\mathcal{A}} I(\mu, Q). \tag{3.5}$$

We point out that Condition $C(\sigma)$ with $\sigma > 0$ implies that $\langle \pi, r\rangle < +\infty$ (cf. [5, Lemma 3.9]) and that $r(\cdot)$ has compact level sets (cf. [5, Remark 2.3]). Moreover, Condition $C(0)$ (i.e. $C(\sigma)$ with $\sigma = 0$) with (i) replaced by the fact that u_n belongs to the domain of the infinitesimal generator of the Markov chain $(\xi_t)_{t\in\mathbb{R}_+}$, and with Lu_n defined as the infinitesimal generator applied to u_n, is the condition under which the large deviation of the empirical measure is derived in [13]–(IV). Finally, see [5, Section 2.3], it holds $I(\mu, Q) = 0$ if and only if $\mu = \pi$ and $Q = Q^\pi$ and Theorem 3.2 implies that, for any $x \in V$, the empirical flow Q_T converges in $L^1_+(E)$ (endowed with the bounded weak* topology) to Q^π in $\mathbb{P}_x$-probability.

Remark 3.3. As discussed in [5] Theorem 3.2 holds also replacing Condition $C(\sigma)$, $\sigma > 0$, with a suitable hypercontractivity assumption (see Condition 2.4 there). Also the results we present in the rest of the present article could be obtained under this alternative assumption.

4. Comments on the main assumptions

We first recall some basic facts from [25, Chapter 3]. Assuming (A1) and irreducibility (A3), assumptions (A2) and (A4) together are equivalent to the fact that all states are positive recurrent. In (A4) one could remove the assumption of uniqueness of the invariant probability measure, since for an irreducible Markov chain there can be at most one. We observe that if V is finite then (A1) and (A2) are automatically satisfied, while (A3) implies (A4).

Proposition 4.1. *If the Markov chain satisfies assumptions* (A1), (A2), (A3) *and Condition $C(\sigma)$ for some $\sigma \ge 0$, then* (A4) *is verified.*

Proof. The core of the proof will consist in showing that there exist $s > 0$ and a probability measure π on V such that $\pi P(s) = \pi$, where $P(s)$ is the $V \times V$-matrix such that $P_{y,z}(s) = \mathbb{P}_y(\xi_s = z)$.

Before proving this property, let us explain how to deduce that π is invariant in the algebraic sense (2.1) (as already stressed, uniqueness in (A4) is a consequence of (A3)). Due to [25, Theorem 3.5.5] we only need to prove that the Markov chain ξ is recurrent. To this aim, consider the discrete time Markov chain $\zeta_n := \xi_{ns}$ with associated stochastic matrix $P(s)$. Note that the irreducibility of ξ implies the irreducibility of ζ and that the condition $\pi P(s) = \pi$ corresponds to the fact that π is an invariant probability for ζ. Hence, due to [25, Theorem 1.7.7], each state is positive recurrent for the Markov chain ζ and therefore is recurrent for the Markov chain ξ.

Next, we exhibit $\pi \in \mathcal{P}(V)$ such that $\pi P(s) = \pi$ for any $s > 0$. To this aim, we fix $x \in V$ and, given an integer $n \geq 1$, we define $\pi_n \in \mathcal{P}(V)$ as $\pi_n(A) = \mathbb{E}_x(\mu_n(A))$ for all $A \subset V$ (μ_n denotes the empirical measure at time n). We claim that, due to Condition $C(\sigma)$, the sequence $\{\pi_n\}_{n\geq 1}$ is tight in $\mathcal{P}(V)$. In the proof of Proposition 3.6 in [5] we have deduced (without using (A4)) that for each $\ell \geq 1$ there exists a finite set $K_\ell \subset V$ such that $\lim_{n\to\infty} \mathbb{P}_x\left(\mu_n(K_\ell^c) > \frac{1}{\ell}\right) = 0$. Since

$$\begin{aligned}\pi_n(K_\ell^c) &= \mathbb{E}_x(\mu_n(K_\ell^c)) \leq \frac{1}{\ell}\mathbb{P}_x\left(\mu_n(K_\ell^c) \leq \frac{1}{\ell}\right) + \mathbb{P}_x\left(\mu_n(K_\ell^c) > \frac{1}{\ell}\right)\\ &\leq \frac{1}{\ell} + \mathbb{P}_x\left(\mu_n(K_\ell^c) > \frac{1}{\ell}\right),\end{aligned}$$

it is simple to obtain that the sequence $\{\pi_n\}_{n\geq 1}$ is tight in $\mathcal{P}(V)$. By Prohorov theorem (cf. [6, Theorem 5.1]) the sequence is relatively compact, and therefore there exists a subsequence $n_k \nearrow \infty$ and a probability measure π in $\mathcal{P}(V)$ such that π_{n_k} converges weakly to π. Let us show that for any $s > 0$ it holds $\pi P(s) = \pi$. To this aim we show that $\langle \pi, P(s) f\rangle = \langle \pi, f\rangle$ for any bounded function $f : V \to \mathbb{R}$. Since $P(s) f : V \to \mathbb{R}$ is bounded and continuous, by the weak convergence we can write

$$\langle \pi, P(s) f\rangle = \lim_{k\to\infty} \langle \pi_{n_k}, P(s) f\rangle. \tag{4.1}$$

On the other hand, given $g : V \to \mathbb{R}$ bounded it holds

$$\langle \pi_n, g\rangle = \mathbb{E}_x\left(\frac{1}{n}\int_0^n g(X_u)du\right) = \frac{1}{n}\int_0^n \mathbb{E}_x\left(g(X_u)\right)du = \frac{1}{n}\int_0^n [P(u)g](x)du.$$

In particular, by using the above identity twice (both with $g := P(s) f$ and with $g := f$) and using the semigroup property $P(u)P(s) = P(u+s)$, we have

$$\begin{aligned}\langle \pi_n, P(s) f\rangle &= \langle \pi_n, f\rangle - \frac{1}{n}\int_0^s [P(u) f](x)du + \frac{1}{n}\int_n^{n+s} [P(u) f](x)du\\ &= \langle \pi_n, f\rangle + O\left(\frac{s}{n}\right).\end{aligned} \tag{4.2}$$

By setting $n := n_k$ in (4.2) and afterwards taking the limit $k \to +\infty$, from the weak convergence of π_{n_k} to π we conclude that (4.1) equals $\langle \pi, f\rangle$. □

5. Joint LDP for the empirical measure and flow in the strong $L^1_+(E)$ topology

As stated in Theorem 2.7.2. in [24], the bounded weak* topology is weaker than the strong topology in $L^1_+(E)$, i.e. the one coming from the L^1-norm. This means that any bounded weakly* open (closed) set is also strongly open (closed).

Proposition 5.1. *Under the same hypotheses of Theorem 3.2 a weak joint LDP for (μ_T, Q_T) holds with the strong topology on $L^1_+(E)$. Namely, (3.4) and (3.5) are valid for any $\mathcal{C}$ compact and any $\mathcal{A}$ open when $L^1_+(E)$ is endowed with the strong topology.*

Proof. Since any strongly compact subset of $L^1_+(E)$ is bounded weak* compact and therefore bounded weak* closed (as the bounded weak* topology is Hausdorff), the upper bound for strongly compact subsets is a direct consequence of (3.4). Moreover, one can verify that the direct proof in [5, Section 5] of the lower bound (3.5) works also for strongly open sets. Note that the hypothesis of locally finite graph stated in [5, Section 5] is not used for the lower bound. □

We now describe a criterion implying the (full) joint LDP for (μ_T, Q_T) when $L^1_+(E)$ is endowed with the strong topology. To this aim, given $E' \subset E$, we define $Q(E') = \sum_{(y,z)\in E'} Q(y, z)$. Moreover, for a fixed subset $\widehat{E} \subset E$, we define the $\widehat{E}$-dependent function $H : V \mapsto \mathbb{R}$ as

$$H(y) := \frac{\sum_{z:(y,z)\in\widehat{E}} r(y,z)}{\sum_{z:(y,z)\in E} r(y,z)}. \tag{5.1}$$

Given $a \in (0, 1)$ suppose that $H(y) < a$. Then, after arriving in y, the Markov chain has probability $H(y) < a$ to jump from y along an edge in $\widehat{E}$. We then call *a-unlikely* all edges (y, z) with $H(y) < a$ and $(y, z) \in \widehat{E}$, while we call *a-likely* all edges (y, z) with $H(y) < a$ and $(y, z) \in E\setminus\widehat{E}$.

Theorem 5.2. *Assume Assumptions* (A1), (A2), (A3) *and Condition $C(\sigma)$ with $\sigma > 0$. Suppose there exists a subset $\widehat{E} \subset E$ such that*

(i) *for each $y \in V$ there exists $z \in V$ with $(y, z) \in \widehat{E}$;*
(ii) *the function $H : V \to (0, +\infty)$ defined in (5.1) vanishes at infinity;*
(iii) *fixed any $x \in V$, there exist constants $a_0, \gamma > 0$ such that for any $a < a_0$ one can find a subset $W = W(x, a)$ in E satisfying the following properties:*
   (1) *the complement $E\setminus W$ is finite;*
   (2) *each edge in W is a-likely or a-unlikely, i.e. if $(y, z) \in W$ then $H(y) < a$;*
   (3) *for each path exiting from x the number of a-unlikely edges in W is at least γ-times the total number of edges in W. Namely, for any path $x_1, x_2, \dots, x_n$ with $x_1 = x$ and $(x_i, x_{i+1}) \in E$ it holds*

$$\sharp\Big\{i : (x_i, x_{i+1}) \in \widehat{E} \cap W\Big\} \geq \gamma \sharp\Big\{i : (x_i, x_{i+1}) \in W\Big\}. \tag{5.2}$$

Then Theorem 3.2 remains valid if $L^1_+(E)$ is endowed with the strong topology instead of the bounded weak topology.*

Applications of the above theorem can be found in Proposition 10.1 and in Lemma 10.4 of Section 10. We point out that a possible natural candidate for the above set W is given by the set $\{(y, z) : H(y) < a\}$. In many applications the geometric control of $\{(y, z) : H(y) < a\}$ is partial, and therefore it can be convenient to use a subset $W \subset \{(y, z) : H(y) < a\}$.

Proof. In view of Proposition 5.1, we only need to prove the exponential tightness of the empirical flow in the strong topology.

The core of the proof consists in showing that there exists an invading sequence of finite subsets $E_n \nearrow E$ such that

$$\mathbb{P}_x\big(Q_T(E_n^c) \geq 1/n\big) \leq c_1 e^{-c_2 T n + c_3 T} \quad \forall T \geq 0,\ \forall n \geq 1, \tag{5.3}$$

for suitable positive constants c_1, c_2, c_3 (depending on $x \in V$). Let us first derive from (5.3) the exponential tightness of the empirical flow in the strong topology. To this aim, fixed positive integers ℓ, m, we let

$$\mathcal{K}_{m,\ell} := \Big\{Q \in L^1_+(E) : \|Q\| \leq \ell,\ Q(E_n^c) \leq 1/n\ \forall n \geq m\Big\}.$$

We claim that $\mathcal{K}_{m,\ell} \subset L^1_+(E)$ is compact for the strong topology. Indeed, by Prohorov theorem for measures [8, Chapter 8], the set $\mathcal{K}_{m,\ell}$ is relatively compact in the space of nonnegative finite measures on E endowed with the weak topology. Hence, given a sequence $\{Q_k\}_{k\geq 0}$ in $\mathcal{K}_{m,\ell}$, at cost to extract a subsequence we can assume that Q_k converges weakly to some $Q : E \to [0, \infty)$ thought of as measure on E. By definition of weak convergence (recall that E has the discrete topology, hence any function on E is continuous) one gets that $Q \in \mathcal{K}_{m,\ell}$ and that $Q_k(e) \to Q(e)$ for all $e \in E$. In particular, one can estimate $\|Q - Q_k\| \leq 2/n + \sum_{e\in E_n} |Q(e) - Q_k(e)|$ for $n \geq m$. This implies that $\|Q - Q_k\|$ converges to 0, hence our claim.

We can bound

$$\mathbb{P}_x(Q_T \notin \mathcal{K}_{m,\ell}) \leq \mathbb{P}_x\big(\|Q_T\| \geq \ell\big) + \sum_{n\geq m} \mathbb{P}_x(Q_T(E_n^c) > 1/n). \tag{5.4}$$

By Proposition 3.6 in [5] $\lim_{\ell\to+\infty} \overline{\lim}_{T\to+\infty} \frac{1}{T} \log \mathbb{P}_x\big(\|Q_T\| \geq \ell\big) = -\infty$, while by (5.3) the series in the above r.h.s. is bounded by $c_1 e^{-c_2 T m + c_3 T}/(1 - e^{-c_2 T})$. This implies the exponential tightness, under $\mathbb{P}_x$, of the empirical flow Q_T in $L^1_+(E)$ endowed with the strong topology.

Let us now derive (5.3). We first point out that, for suitable positive constants λ and c, it holds

$$\mathbb{E}_x\Big\{e^{T\lambda\langle\mu_T, r\rangle}\Big\} \leq c\, e^{cT}, \quad \forall T \geq 0. \tag{5.5}$$

Indeed, this follows from [5, Lemma 3.5] with $\lambda = \sigma$ (if instead of Condition $C(\sigma)$ one assumes Items (i) and (ii) of the hypercontractivity Condition 2.4 in [5], then (5.5) follows from [5, Eq. (3.12)]).

Fixed λ as above, we introduce the set $\widetilde{E} := \{(y, z) \in \widehat{E} : H(y) \leq \lambda\}$ and then define the function $F : E \to [0, +\infty)$ as

$$F(y, z) := \begin{cases} \log \dfrac{\lambda}{H(y)} & \text{if } (y, z) \in \widetilde{E}, \\ 0 & \text{if } (y, z) \in E\setminus\widetilde{E}. \end{cases}$$

Defining $r^F(y, z) := r(y, z)e^{F(y,z)}$ we get (recall that $r(y) = \sum_{z:(y,z)\in E} r(y, z)$):

$$\begin{aligned} r^F(y) := \sum_{z:(y,z)\in E} r^F(y, z) &= \sum_{z:(y,z)\in E\setminus\widetilde{E}} r(y, z) + \frac{\lambda}{H(y)} \sum_{z:(y,z)\in\widetilde{E}} r(y, z) \\ &\leq \sum_{z:(y,z)\in E} r(y, z) + \frac{\lambda}{H(y)} \sum_{z:(y,z)\in\widehat{E}} r(y, z) \leq (1 + \lambda)r(y). \end{aligned} \tag{5.6}$$

In particular, we conclude that $r^F(y) - r(y) \le \lambda r(y)$ for all $y \in V$. Since $r^F(y) < +\infty$ for all $y \in V$, by Lemma 3.1 in [5] we get that

$$\mathbb{E}_x\left\{e^{T\left[\langle Q_T,F\rangle - \lambda\langle \mu_T, r\rangle\right]}\right\} \le \mathbb{E}_x\left\{e^{T\left[\langle Q_T,F\rangle - \langle \mu_T, r^F - r\rangle\right]}\right\} \le 1. \tag{5.7}$$

By Schwarz inequality, combining (5.5) and (5.7), we get for some $C > 0$:

$$\mathbb{E}_x\left\{e^{\frac{T}{2}\langle Q_T,F\rangle}\right\} \le \mathbb{E}_x\left\{e^{T\left[\langle Q_T,F\rangle - \lambda\langle \mu_T, r\rangle\right]}\right\}^{\frac{1}{2}} \mathbb{E}_x\left\{e^{T\lambda\langle \mu_T, r\rangle}\right\}^{\frac{1}{2}} \le Ce^{CT}. \tag{5.8}$$

Take $a < a_0$ and recall the properties of $W(x, a) \subset E$ given in Item (iii). Since x is fixed, we write simply $W(a)$. By assumption $W(a)^c$ is a finite set. Given an integer $n \ge 1$ let $a_n := \lambda/e^{n^2}$. In particular if $H(y) \le a_n$ it must be $H(y) < \lambda$ and $\ln(\lambda/H(y)) \ge n^2$. We conclude that

$$F(y, z) \ge n^2 \quad \forall (y, z) \in \widetilde{E} \cap W(a_n) = \widehat{E} \cap W(a_n). \tag{5.9}$$

Since F is a nonnegative function, combining (5.8) with (5.9) we get

$$\mathbb{E}_x\left\{e^{\frac{n^2}{2} T Q_T(\widehat{E}\cap W(a_n))}\right\} \le Ce^{CT}. \tag{5.10}$$

On the other hand, by applying Item (iii)-(3) to the family of consecutive states visited by the trajectory $(X_t)_{t\in[0,T]}$, we get that $T Q_T(\widehat{E} \cap W(a_n)) \ge \gamma T Q_T\big(W(a_n)\big)$. Hence we conclude that

$$\mathbb{E}_x\left\{e^{\frac{n^2\gamma}{2} T Q_T(W(a_n))}\right\} \le Ce^{CT}. \tag{5.11}$$

Consider now the set $E_n := W(a_n)^c$, which is finite by Item (iii)-(1). By Chebyshev inequality and (5.11) we obtain

$$\mathbb{P}_x\left(Q_T(E_n^c) \ge \frac{1}{n}\right) = \mathbb{P}_x\left(\frac{n^2\gamma}{2} T Q_T(W(a_n)) \ge \frac{n\gamma T}{2}\right) \le Ce^{-\frac{n\gamma T}{2} + CT},$$

thus leading to (5.3). □

6. Joint large deviations for the empirical measure and current

Recalling that E denotes the set of ordered edges in V with strictly positive jump rate, we let $E_{\mathrm{s}} := \big\{(y, z) \in V \times V : (y, z) \in E \text{ or } (z, y) \in E\big\}$ be the symmetrization of E in $V \times V$. We then introduce $L^1_{\mathrm{a}}(E_{\mathrm{s}})$ as the space of antisymmetric and absolutely summable functions on E_{s}, i.e.

$$L^1_{\mathrm{a}}(E_{\mathrm{s}}) := \big\{J \in L^1(E_{\mathrm{s}}) : J(y, z) = -J(z, y)\ \forall\,(y, z) \in E_{\mathrm{s}}\big\}.$$

Elements of $L^1_{\mathrm{a}}(E_{\mathrm{s}})$ will be denoted by J and called *currents*. We shall consider $L^1_{\mathrm{a}}(E_{\mathrm{s}})$ endowed either with the bounded weak* topology or with the strong topology, and the associated Borel σ-algebra.

To each flow $Q \in L^1_+(E)$, we associate the *canonical current* J_Q defined by

$$J_Q(y, z) := \begin{cases} Q(y, z) - Q(z, y) & \text{if } (y, z) \in E \text{ and } (z, y) \in E, \\ Q(y, z) & \text{if } (y, z) \in E \text{ and } (z, y) \notin E, \\ -Q(z, y) & \text{if } (y, z) \notin E \text{ and } (z, y) \in E. \end{cases} \tag{6.1}$$

Given a current J we define its divergence, $\mathrm{div} J \in L^1(V)$ by

$$\mathrm{div} J(y) := \sum_{z:(y,z)\in E_\mathrm{s}} J(y,z).$$

It is simple to check the above definition is consistent with (2.4) in the sense that $\mathrm{div} J_Q = \mathrm{div} Q$.

Given $T > 0$, the *empirical current* is the map $J_T : D(\mathbb{R}_+; V) \to L^1_\mathrm{a}(E_\mathrm{s})$ defined as

$$\begin{aligned} J_T(y,z)(X) &:= \frac{1}{T} \sum_{0\le t\le T} \left[\mathbb{1}\big(X_{t^-} = y,\ X_t = z\big) - \mathbb{1}\big(X_{t^-} = z, X_t = y\big)\right] \\ &= Q_T(y,z)(X) - Q_T(z,y)(X) \end{aligned} \tag{6.2}$$

for all $(y,z) \in E_s$, where $Q_T(a,b) := 0$ if $(a,b) \in E_s \setminus E$. Namely, $T J_T(y,z)$ is the *net* number of jumps across $(y,z) \in E_\mathrm{s}$ in the time interval $[0,T]$. Equivalently, the empirical current J_T is the canonical current associated to the empirical flow Q_T, i.e. $J_T = J_{Q_T}$.

Recalling (2.5), to each probability $\mu \in \mathcal{P}(V)$ we associate the current $J^\mu := J_{Q^\mu}$, i.e. $J^\mu(y,z) = \mu(y)r(y,z) - \mu(z)r(z,y)$. Observe that J^μ has vanishing divergence if and only if $\mu = \pi$ and J^μ vanishes if and only if the chain is reversible with respect to μ. In view of the discussion in Section 2.1, for each $x \in V$ and $(y,z) \in E_\mathrm{s}$ the sequence of real random variables $\{J_T(y,z)\}$ converges, in probability with respect to $\mathbb{P}_x$, to $J^\pi(y,z)$ as $T \to +\infty$.

To state the joint LDP for (μ_T, J_T) we introduce the function $\Psi : \mathbb{R}\times\mathbb{R}\times\mathbb{R}_+ \mapsto [0,+\infty)$ given by

$$\Psi(u,\bar u; a) := \begin{cases} u\left[\operatorname{arcsinh}\dfrac{u}{a} - \operatorname{arcsinh}\dfrac{\bar u}{a}\right] - \left[\sqrt{a^2+u^2} - \sqrt{a^2+\bar u^2}\right] & \text{if } a > 0, \\ \Phi(u,\bar u) & \text{if } a = 0. \end{cases} \tag{6.3}$$

Due to the continuity of the map $Q \mapsto J_Q$ the joint large deviation principle for the empirical measure and current follows from Theorem 3.2 by contraction:

Theorem 6.1. *Assume the Markov chain satisfies* (A1), (A2), (A3) *and Condition $C(\sigma)$ with $\sigma > 0$. Then, as $T \to +\infty$, the sequence of probability measures $\{\mathbb{P}_x \circ (\mu_T, J_T)^{-1}\}$ on $\mathcal{P}(V) \times L^1_\mathrm{a}(E_\mathrm{s})$ satisfies a large deviation principle with good and convex rate function $\widetilde{I} : \mathcal{P}(V) \times L^1_\mathrm{a}(E_\mathrm{s}) \to [0,+\infty]$.*

To have $\widetilde{I}(\mu, J) < +\infty$ it is necessary that $\mathrm{div} J = 0$, $J(y,z) \ge 0$ for any $(y,z) \in E$ such that $(z,y) \notin E$ and $\langle \mu, r\rangle < +\infty$. When all these conditions are satisfied we have

$$\widetilde{I}(\mu,J) = I(\mu, Q^{J,\mu}) = \sum_{(y,z)\in E} \Phi\big(Q^{J,\mu}(y,z), Q^\mu(y,z)\big), \tag{6.4}$$

where

$$Q^{J,\mu}(y,z) := \frac{J(y,z) + \sqrt{J^2(y,z) + 4\mu(y)\mu(z)r(y,z)r(z,y)}}{2}.$$

The above identity (6.4) can also be rewritten as

$$\widetilde{I}(\mu,J) = \frac{1}{2} \sum_{(y,z)\in E_s} \Psi\big(J(y,z), J^\mu(y,z); a^\mu(y,z)\big) \tag{6.5}$$

where

$$a^\mu(y,z) := 2\sqrt{\mu(y)\mu(z)r(y,z)r(z,y)}.$$

Moreover, if the conditions of Theorem 5.2 are satisfied, then the above result remains true with $L^1_{\rm a}(E_{\rm s})$ endowed with the strong L^1-topology.

Note that if $J \in L^1_{\rm a}(E_{\rm s})$ and $\langle \mu, r\rangle < +\infty$ then $Q^{J,\mu} \in L^1_+(E)$, moreover if $\mathrm{div} J = 0$ then also $\mathrm{div} Q^{J,\mu} = 0$. Note also that $\widetilde{I}(\mu, J) = 0$ if and only if $I(\mu, Q^{J,\mu}) = 0$, and we know this holds if and only if $\mu = \pi$ and $Q^{J,\mu} = Q^\pi$ (see the discussion after Theorem 3.2). It is trivial to check that this last condition is equivalent to $(\mu, J) = (\pi, J^\pi)$.

Proof of Theorem 6.1. Recalling that $L^1_+(E)$ is equipped with the bounded weak* topology, the map $L^1_+(E) \ni Q \mapsto J_Q \in L^1_{\rm a}(E_{\rm s})$ is continuous. This map remains continuous if $L^1_+(E)$ and $L^1_{\rm a}(E_{\rm s})$ are both endowed of the strong L^1-topology. Hence, by Theorem 3.2, Theorem 5.2 and the contraction principle, a joint LDP holds for (μ_T, J_T) with good rate function

$$\widetilde{I}(\mu, J) := \inf\big\{ I(\mu, Q) : Q \in L^1_+(E) \text{ with } J_Q = J \big\}, \quad (\mu, J) \in \mathcal{P}(V) \times L^1_a(E_s).$$

It remains to show that the above $\widetilde{I}(\mu, J)$ fulfills the properties stated in the theorem.

From the above variational characterization of $\widetilde{I}(\mu, J)$ one easily derives that $\widetilde{I}$ is convex, since $I(\mu, Q)$ is convex and the map $L^1_+(E) \ni Q \to J_Q \in L^1_a(E_s)$ is linear.

It is simple to verify that to have $\widetilde{I}(\mu, J) < +\infty$ it is necessary that $\mathrm{div} J = 0$, $J(y, z) \geq 0$ for any $(y, z) \in E$ such that $(z, y) \notin E$ and $\langle \mu, r\rangle < +\infty$. Let us now take $(\mu, J) \in \mathcal{P}(V) \times L^1_{\rm a}(E_{\rm s})$ satisfying the above three conditions. Then the set $\big\{ Q \in L^1_+(E) : J_Q = J \big\}$ coincides with the set of flows of the type

$$Q(y, z) = \begin{cases} [J(y, z)]_+ + s(\{y, z\}), & \text{if } (z, y) \in E, \\ [J(y, z)]_+ = J(y, z), & \text{if } (z, y) \notin E, \end{cases}$$

where $[\cdot]_+$ denotes the positive part (i.e. $[z]_+ := \max\{0, z\}$), $s \in L^1_+(E_{\rm u})$ and $E_{\rm u} := \{\{y, z\} : (y, z) \in E_{\rm s}\}$ is the set of *unordered edges*. We can then solve independently a variational problem for each pair of edges (y, z) and (z, y) in $E_{\rm s}$. If (y, z) and (z, y) both belong to E, then an elementary computation gives that

$$\begin{aligned} &\inf_{s \in [0, +\infty)} \Big\{ \Phi\Big([J(y, z)]_+ + s, Q^\mu(y, z) \Big) + \Phi\Big([-J(y, z)]_+ + s, Q^\mu(z, y) \Big) \Big\} \\ &\quad = \Phi\Big(Q^{J,\mu}(y, z), Q^\mu(y, z) \Big) + \Phi\Big(Q^{J,\mu}(z, y), Q^\mu(z, y) \Big). \end{aligned}$$

If $(y, z) \in E$ and $(z, y) \notin E$, then $Q(y, z) = J(y, z) = Q^{J,\mu}(y, z)$ for any $Q \in L^1_+(E)$ with $J_Q = J$. Since $\mathrm{div} Q = \mathrm{div} J_Q = 0$, by the expression (3.3) of the rate function I and the above computations, we obtain that $\widetilde{I}(\mu, Q) = I(\mu, Q^{J,\mu})$, hence (6.4).

It remains to prove (6.5). To this aim we observe that, if both (y, z) and (z, y) belong to E, then the following identities hold:

$$\begin{aligned} &\Phi\big(Q^{J,\mu}(y, z), Q^\mu(y, z) \big) + \Phi\big(Q^{J,\mu}(z, y), Q^\mu(z, y) \big) \\ &\quad = \frac{j}{2} \log\Big[\frac{j + \sqrt{j^2 + 4pp'}}{-j + \sqrt{j^2 + 4pp'}} \frac{p'}{p} \Big] - \sqrt{j^2 + 4pp'} + p + p' \\ &\quad = j \log \frac{j + \sqrt{j^2 + 4pp'}}{2p} - \sqrt{j^2 + 4pp'} + p + p' \end{aligned} \tag{6.6}$$

where $j := J(y, z)$, $p = \mu(y) r(y, z)$, $p' = \mu(z) r(z, y)$, assuming p, p' positive. Set $a := a^\mu(y, z) = 2\sqrt{pp'}$ and $\bar{j} := J^\mu(y, z) = p - p'$. Since $\mathrm{arcsinh}\, u = \log[u + \sqrt{u^2 + 1}]$, $j^2 + 4pp' = j^2 + a^2$, $p + p' = \sqrt{\bar{j}^2 + a^2}$, the last member in (6.6) can be rewritten as $\Psi(j, \bar{j}; a)$.

Suppose now that $(y, z) \in E$ and $(y, z) \notin E$, $J(y, z) \geq 0$. Then $Q^{J,\mu}(y, z) = J(y, z)$ and $Q^{\mu}(y, z) = J^{\mu}(y, z)$. In particular,

$$\Phi\big(Q^{J,\mu}(y, z), Q^{\mu}(y, z)\big) = \Psi(J(y, z), J^{\mu}(y, z); 0).$$

From the above considerations it is simple to derive (6.5) from (6.4). □

7. Gallavotti–Cohen type symmetries for the empirical current

In this section and in Sections 8 and 9, we assume that $E_s = E$ (i.e. $r(y, z) > 0$ if and only if $r(z, y) > 0$) and we derive Gallavotti–Cohen (GC) symmetries of the LD rate function both of the empirical current and of suitable linear functionals of the empirical current itself.

In what follows, $w_\pi : E \to \mathbb{R}$ denotes the antisymmetric function

$$w_\pi(y, z) = \log \frac{\pi(y)r(y, z)}{\pi(z)r(z, y)}, \quad (y, z) \in E, \tag{7.1}$$

and we will assume that $w_\pi \in L^\infty(E)$, thus implying that $\langle J, w_\pi\rangle$ is finite for any $J \in L^1_a(E)$.

Theorem 7.1. *Assume $E_s = E$ and that $w_\pi \in L^\infty(E)$. Then the rate function $\widetilde{I}$ of Theorem 6.1 satisfies the following GC symmetry in $[0, +\infty]$:*

$$\widetilde{I}(\mu, J) = \widetilde{I}(\mu, -J) - \frac{1}{2}\langle J, w_\pi\rangle, \quad \forall(\mu, J) \in \mathcal{P}(V) \times L^1_a(E). \tag{7.2}$$

In particular, the good and convex rate function $\widehat{I} : L^1_a(E) \to [0, +\infty]$, $\widehat{I}(J) = \inf_\mu \widetilde{I}(\mu, J)$, of the LDP for the empirical current obtained by contraction from Theorem 6.1 satisfies the following GC symmetry in $[0, +\infty]$:

$$\widehat{I}(J) = \widehat{I}(-J) - \frac{1}{2}\langle J, w_\pi\rangle, \quad \forall J \in L^1_a(E). \tag{7.3}$$

Remark 7.2. For finite state spaces the GC symmetry (7.3) has already been derived in [1,2,14] in terms of the moment generating functions (essentially, by means of Gärtner–Ellis theorem).

Proof. Having (7.2), the conclusion is a trivial consequence of the contraction principle. Let us prove (7.2). If $\operatorname{div} J \neq 0$ or $\langle \mu, r\rangle = +\infty$, then $\widetilde{I}(\mu, J) = \widetilde{I}(\mu, -J) = +\infty$ and (7.2) is trivially true (recall that $\langle J, w_\pi\rangle$ is finite). Suppose therefore that $\operatorname{div} J = 0$ and $\langle \mu, r\rangle < +\infty$. Then, $\widetilde{I}(\mu, J)$ and $\widetilde{I}(\mu, -J)$ have the series expression induced by (6.4). It is simple to check that, given $p, p' > 0$ and $q, q' \geq 0$, it holds

$$\Phi(q, p) + \Phi(q', p') = \Phi(q', p) + \Phi(q, p') + (q - q')\log(p'/p). \tag{7.4}$$

Taking $q := Q^{J,\mu}(y, z) = Q^{-J,\mu}(z, y)$, $q' := Q^{J,\mu}(z, y) = Q^{-J,\mu}(y, z)$, $p := Q^{\mu}(y, z)$, $p' := Q^{\mu}(z, y)$, from the above identity we get

$$\begin{aligned}&\Phi\Big(Q^{J,\mu}(y, z), Q^{\mu}(y, z)\Big) + \Phi\Big(Q^{J,\mu}(z, y), Q^{\mu}(z, y)\Big)\\ &\quad = \Phi\Big(Q^{-J,\mu}(y, z), Q^{\mu}(y, z)\Big) + \Phi\Big(Q^{-J,\mu}(z, y), Q^{\mu}(z, y)\Big) - J(y, z)w_\pi(y, z).\end{aligned}$$

Summing above $(y, z) \in E$ we get (7.2). □

8. Gallavotti–Cohen symmetry for the Gallavotti–Cohen functional

Let $\mathbb{P}_\pi$ be the law of the stationary chain (the initial state is sampled according to the invariant probability π). By stationarity, $\mathbb{P}_\pi$ can be extended to a measure on $D(\mathbb{R}; V)$. Let $\vartheta : D(\mathbb{R}; V) \to D(\mathbb{R}; V)$ be the time reversal, i.e. for the set of times $t \in \mathbb{R}$ which are continuity points of X the map ϑ is defined by $(\vartheta X)_t = X_{-t}$. We then set $\mathbb{P}^*_\pi := \mathbb{P}_\pi \circ \vartheta^{-1}$; of course $\mathbb{P}^*_\pi = \mathbb{P}_\pi$ if and only if the chain is reversible. In general, $\mathbb{P}^*_\pi$ is the law of the stationary chain with jump rates $r^*(y, z) = \pi(z)r(z, y)/\pi(y)$. Given $x \in V$ and $T > 0$, the Gallavotti–Cohen functional can be defined (cf. [20]) as the map $W_T : D(\mathbb{R}_+; V) \to \mathbb{R}$ which is $\mathbb{P}_x$ a.s. given by

$$W_T := -\frac{1}{T} \log \frac{d\mathbb{P}^*_\pi\big|_{[0,T]}}{d\mathbb{P}_\pi\big|_{[0,T]}}. \tag{8.1}$$

Observe that $\mathbb{E}_\pi\big(W_T\big)$ is $(1/T)$-proportional to the relative entropy of $\mathbb{P}_\pi|_{[0,T]}$ with respect to $\mathbb{P}^*_\pi\big|_{[0,T]}$, thus providing a natural measure of the irreversibility of the chain.

A simple computation of the Radon–Nikodym derivative in (8.1) (use (3.1) in [5] and observe that $r^*(y) = r(y)$ for any $y \in V$ due to the invariance of π) gives that the Gallavotti–Cohen functional W_T can be written in terms of the empirical current J_T as

$$W_T = \frac{1}{2} \langle J_T, w_\pi \rangle, \tag{8.2}$$

where $w_\pi : E \to \mathbb{R}$ is the antisymmetric function defined by (7.1).

As a consequence of our previous results and the contraction principle we get the following LDP:

Theorem 8.1. *Assume that $E = E_s$, the Markov chain satisfies* (A1), (A2), (A3) *and assume Condition $C(\sigma)$ with $\sigma > 0$. Assume also that w_π vanishes at infinity. If the conditions of Theorem 5.2 are satisfied, it is enough to require that w_π is a bounded function.*

Then, as $T \to +\infty$, the sequence of probability measures $\{\mathbb{P}_x \circ W_T^{-1}\}$ on $\mathbb{R}$ satisfies a large deviation principle with good and convex rate function $\imath : \mathbb{R} \to [0, +\infty]$ given by

$$\imath(u) = \inf\big\{\widetilde{I}(\mu, J) : (\mu, J) \in \mathcal{P}(V) \times L^1_{\mathrm{a}}(E),\ \langle J, w_\pi\rangle = 2u\big\}. \tag{8.3}$$

Moreover, the following GC symmetry holds in $[0, +\infty]$*:*

$$\imath(u) = \imath(-u) - u. \tag{8.4}$$

Proof. Note that the map $L^1_{\mathrm{a}}(E) \ni J \to \langle J, w_\pi\rangle \in \mathbb{R}$ is well defined and continuous in both the following cases: (i) $L^1_{\mathrm{a}}(E)$ is endowed of the bounded weak* topology and w_π vanishes at infinity, (ii) $L^1_{\mathrm{a}}(E)$ is endowed of the strong L^1-topology and w_π is bounded. Hence, due to the contraction principle and Theorem 6.1, we only need to prove that the rate function $\imath(\cdot)$ is convex and that GC symmetry (8.4) is fulfilled. The last property follows from Theorem 7.1. The convexity follows easily from the fact that $\widetilde{I}(\mu, J)$ is convex and the constraint $\langle J, w_\pi\rangle = 2u$ is linear in J. □

The Gallavotti–Cohen functional is defined in [20] by replacing the function w_π above with $w(y, z) = \log[r(y, z)/r(z, y)]$. In order to be able to discuss applications to Markov chains with infinitely many states we have chosen the previous definition with w_π. Note that

$$\frac{1}{2}\langle J_T, w_\pi\rangle - \frac{1}{2}\langle J_T, w\rangle = \frac{1}{2} \sum_{(y,z)\in E} J_T(y, z) \log \frac{\pi(y)}{\pi(z)} = \frac{1}{T} \log \frac{\pi(X_T)}{\pi(X_0)}. \tag{8.5}$$

Hence, if V is finite, the term $\log \frac{\pi(X_T)}{\pi(X_0)}$ is bounded, thus implying that $\frac{1}{2}\langle J_T, w_\pi\rangle$ and $\frac{1}{2}\langle J_T, w\rangle$ satisfy the same LDP. Theorem 8.1 provides a variational characterization of the rate function for the Gallavotti–Cohen functional which can be compared to the rather implicit one derived e.g. in [20, with w instead of w_π] by using the Perron–Frobenius and the Gärtner–Ellis theorems.

9. LDP for the homological coefficients and Gallavotti–Cohen symmetry

Also in this section we assume that the (connected) graph $G = (E, V)$ has the property $(y, z) \in E \Leftrightarrow (z, y) \in E$ (i.e. $E = E_s$) and extend to the infinite case the concept of *cycle space*. We refer e.g. to [1,2,14,27] for physical applications and e.g. to [9,12] for a mathematical treatment in finite graphs. We also prove that the cycle space is isomorphic to the first cellular homological class over $\mathbb{R}$ of the graph G (shortly, $H_1(G, \mathbb{R})$). Then we associate to each trajectory up to time T a cycle $\mathcal{C}_T$ and prove a LDP for the empirical homological coefficients, which are given by the coefficients in a given basis of the cycle $\mathcal{C}_T$ thought of as element of the cycle space, and therefore of $H_1(G, \mathbb{R})$.

9.1. *Cycle space of the graph G*

We point out that, working with a graph $G = (V, E)$ with $E = E_s$, all information encoded in G corresponds to the one encoded in its unoriented version $G_u = (V, E_u)$, where $E_u := \big\{\{y, z\} : (y, z) \in E\big\}$. The subscript "u" stays for *unoriented*. Hence, the discussion that follows applies as well to unoriented graphs.

We fix some notation. Given an edge $e = (y, z) \in E$ we write $\bar{e} = (z, y)$ for the reversed edge. A *cycle* $\mathcal{C}$ in G is a finite string $(x_1, \dots, x_k)$ of elements of V such that $(x_i, x_{i+1}) \in E$ when $i = 1, \dots, k$, with the convention that $x_{k+1} = x_1$. Given a cycle $\mathcal{C}$ and given $e \in E$ we definite $S_e(\mathcal{C})$ as the number of times the edge e appears in $\mathcal{C}$ minus the number of times the reversed edge $\bar{e}$ appears in $\mathcal{C}$:

$$S_e(\mathcal{C}) := \sharp\{i : 1 \le i \le k,\ (x_i, x_{i+1}) = e\} - \sharp\{i : 1 \le i \le k,\ (x_i, x_{i+1}) = \bar{e}\}.$$

Consider now the free real vector space $\mathcal{V}$ generated by all cycles $\mathcal{C}$. Its elements are the formal sums $\sum_{j=1}^n a_j \mathcal{C}_j$, varying $n \in \mathbb{N}$, $a_j \in \mathbb{R}$ and $\mathcal{C}_j$ cycles, with the natural rules for sum and multiplication by a constant. The empty sum is the zero element of $\mathcal{V}$, denoted by $\emptyset$.

The *cycle space* $\mathcal{V}_*$ of the graph G is then defined as the quotient vector space of $\mathcal{V}$ imposing that in $\mathcal{V}_*$ it holds

$$\sum_{j=1}^n a_j \mathcal{C}_j = \sum_{i=1}^m b_i \mathcal{C}'_i \quad \text{iff} \quad \sum_{j=1}^n a_j S_e(\mathcal{C}_j) = \sum_{i=1}^m b_i S_e(\mathcal{C}'_i)\ \forall e \in E \tag{9.1}$$

(we keep the same notation for the elements of $\mathcal{V}$ and $\mathcal{V}_*$). More precisely, $\mathcal{V}_*$ is defined as the quotient $\mathcal{V}/\mathcal{W}$, where the subspace $\mathcal{W}$ is given by the sums $\sum_{j=1}^n a_j \mathcal{C}_j - \sum_{i=1}^m b_i \mathcal{C}'_i$ satisfying the identity system in the r.h.s. of (9.1). Note that in the cycle space $\mathcal{V}_*$ the cycle $\mathcal{C} = (x_1, x_2, \dots, x_k)$ equals the cycle $\mathcal{C}' = (x_i, x_{i+1}, \dots, x_k, x_1, \dots, x_{i-1})$, and that $-\mathcal{C} = (x_k, x_{k-1}, \dots, x_1)$.

Special bases (called *fundamental bases*) of $\mathcal{V}_*$ can be obtained starting from a spanning tree $\mathcal{T} = (V, E_\mathcal{T})$ of the unoriented graph $G_u = (V, E_u)$. Fix such a spanning tree $\mathcal{T}$. To each edge in $E_u \setminus E_\mathcal{T}$ we assign an orientation and we call *chords* the resulting oriented edges.[1] To each

[1] Usually, chords are the unoriented edges in $E_u \setminus E_\mathcal{T}$ [9,12,27]. To avoid additional notation we have directly included in their definition a fixed orientation.

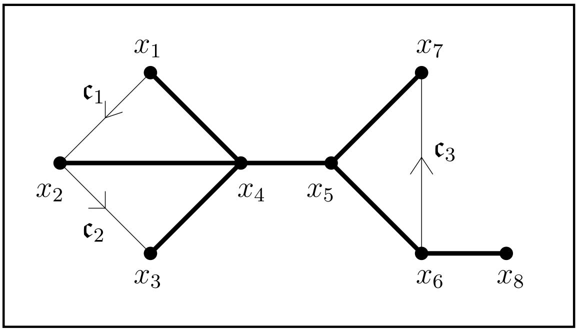

Fig. 1. Fundamental basis. The bold edges form the spanning tree $\mathcal{T}$. The chords $\mathfrak{c}_1 = (x_1, x_2)$, $\mathfrak{c}_2 = (x_2, x_3)$ and $\mathfrak{c}_3 = (x_6, x_7)$ correspond to the cycles $\mathcal{C}_{\mathfrak{c}_1} = (x_1, x_2, x_4)$, $\mathcal{C}_{\mathfrak{c}_2} = (x_2, x_3, x_4)$ and $\mathcal{C}_{\mathfrak{c}_3} = (x_6, x_7, x_5)$, respectively.

chord $\mathfrak{c}$ we associate a cycle $\mathcal{C}_{\mathfrak{c}} \in \mathcal{V}_*$ as follows: writing $\mathfrak{c} = (x_1, x_2)$ consider the unique self-avoiding path $x_2, x_3, \dots, x_{k-1}, x_k, x_1$ from x_2 to x_1 in G such that $\{x_i, x_{i+1}\} \in E_{\mathcal{T}}$ for all $i = 2, 3, \dots, k$ (where $x_{k+1} := x_1$), and set $\mathcal{C}_{\mathfrak{c}} := (x_1, x_2, \dots, x_k)$. Note that by construction

$$S_{\mathfrak{c}}(\mathcal{C}_{\mathfrak{c}'}) = \delta_{\mathfrak{c},\mathfrak{c}'}. \tag{9.2}$$

Proposition 9.1. *Given a spanning tree $\mathcal{T}$ of the unoriented graph $G_u = (V, E_u)$, the cycles $\mathcal{C}_{\mathfrak{c}}$– with $\mathfrak{c}$ varying among the chords of $\mathcal{T}$– form a basis of the quotient space $\mathcal{V}_*$. Moreover, for each cycle $\mathcal{C}$ the following identity holds in $\mathcal{V}_*$:*

$$\mathcal{C} = \sum_{\mathfrak{c}} S_{\vec{\mathfrak{c}}}(\mathcal{C})\mathcal{C}_{\mathfrak{c}}. \tag{9.3}$$

We call the above basis $\{\mathcal{C}_{\mathfrak{c}} : \mathfrak{c} \text{ chord of } \mathcal{T}\}$ a fundamental basis associated to the spanning tree $\mathcal{T}$ (see Fig. 1). Not all basis of $\mathcal{V}_*$ are fundamental, as can be seen e.g. from the simple example given in [14, Section 7]. Due to the above proposition and (9.2), the linear functions $\mathcal{V}_* \ni \sum_{i=1}^n a_i \mathcal{C}_i \mapsto \sum_{i=1}^n a_i S_{\mathfrak{c}}(\mathcal{C}_i) \in \mathbb{R}$, as $\mathfrak{c}$ varies among the chords, form the dual basis of $\{\mathcal{C}_{\mathfrak{c}} : \mathfrak{c} \text{ chord of } \mathcal{T}\}$.

The above proposition is a classical result in the finite setting (cf. [12] when working with the field $\mathbb{F}_2$ instead of $\mathbb{R}$). The proof for infinite graphs could be recovered by the result for finite graphs. For completeness we give a direct and self-contained proof.

Proof. If $\mathcal{C} = \sum_{i=1}^n a_i \mathcal{C}_{\mathfrak{c}_i}$ with chords $\mathfrak{c}_1, \mathfrak{c}_2, \dots, \mathfrak{c}_n$ all distinct, by applying $S_{\mathfrak{c}_j}$ and invoking (9.2) we get that $a_j = S_{\mathfrak{c}_j}(\mathcal{C})$. This proves (9.3) for any cycle $\mathcal{C}$ that is generated by the *fundamental cycles* $\mathcal{C}_{\mathfrak{c}}$'s. We thus need to prove that these cycles form a basis.

We first prove that the cycles $\mathcal{C}_{\mathfrak{c}}$'s are linearly independent. Suppose that $\sum_{i=1}^n a_i \mathcal{C}_{\mathfrak{c}_i} = 0$ for some constants $a_1, \dots, a_n$ and some chords $\mathfrak{c}_1, \dots, \mathfrak{c}_n$. By (9.1) and (9.2) one easily gets that $a_i = 0$ for all i, hence the independence.

We now prove that the cycles $\mathcal{C}_{\mathfrak{c}}$'s generate all $\mathcal{V}_*$. To this end, it is enough to show that they generate any cycle $\mathcal{C}$. Since any cycle $\mathcal{C}$ is in $\mathcal{V}_*$ the sum of self-avoiding cycles, we can restrict to a self-avoiding cycle $\mathcal{C}$, i.e. $\mathcal{C} = (x_1, x_2, \dots, x_k)$ with $x_1, x_2, \dots, x_k$ all distinct (recall that it must be $(x_i, x_{i+1}) \in E$ for all $i = 1, 2, \dots, k$ with the convention $x_{k+1} = x_k$). We prove that the self-avoiding cycle $\mathcal{C}$ can be expressed as linear combination of the fundamental cycles $\mathcal{C}_{\mathfrak{c}}$'s by induction on the cardinality of the set

$$\big\{\mathfrak{c} \text{ chord} : S_{\mathfrak{c}}(\mathcal{C}) \neq 0\big\}. \tag{9.4}$$

If the above set has zero cardinality, i.e. it is empty, then, as $\mathcal{C}$ is self-avoiding and $\mathcal{T}$ is a tree, then $\mathcal{C} = (x_1, x_2)$, which is indeed zero in $\mathcal{V}_*$. Given a positive integer m, let us now suppose that $\mathcal{C}$ is generated by fundamental cycles when the set (9.4) has cardinality less then m. Take a self-avoiding cycle $\mathcal{C} = (x_1, x_2, \dots, x_k)$ such that the set (9.4) has cardinality m and fix a chord $\mathfrak{c}_*$ inside (9.4). Without restriction we can suppose that $(x_2, x_1) = \mathfrak{c}_*$ (at cost to replace $\mathcal{C}$ by $-\mathcal{C}$ and to relabel the points $x_1, x_2, \dots, x_n$). The cycle $\mathcal{C}_{\mathfrak{c}_*}$ is then of the form $(x_2, x_1, y_3, \dots, y_r)$, where $\{x_1, y_3\}, \{y_3, y_4\}, \dots, \{y_{r-1}, y_r\}, \{y_r, x_2\}$ are edges of the tree $\mathcal{T}$. Consider now the cycle

$$\bar{\mathcal{C}} := (x_1, y_3, \dots, y_r, x_2, x_3, \dots, x_m),$$

obtained by removing from $\mathcal{C}$ the edge (x_1, x_2) and replacing it with the path $x_1, y_3, \dots, y_r$. Note that $\mathcal{C} = \bar{\mathcal{C}} - \mathcal{C}_{\mathfrak{c}_*}$ in $\mathcal{V}_*$. By construction,

$$\left\{\mathfrak{c} \text{ chord} : S_{\mathfrak{c}}(\bar{\mathcal{C}}) \neq 0\right\} = \left\{\mathfrak{c} \text{ chord} : S_{\mathfrak{c}}(\mathcal{C}) \neq 0\right\} \setminus \{\mathfrak{c}_*\}. \tag{9.5}$$

At this point, write $\bar{\mathcal{C}}$ as sum $\sum_{u=1}^{s} \bar{\mathcal{C}}_u$ of self-avoiding cycles simply by cutting $\bar{\mathcal{C}}$ at its intersection points. Since the support of $\bar{\mathcal{C}}_u$ is included in the support of $\bar{\mathcal{C}}$ we have

$$\left\{\mathfrak{c} \text{ chord} : S_{\mathfrak{c}}(\bar{\mathcal{C}}_u) \neq 0\right\} \subset \left\{\mathfrak{c} \text{ chord} : S_{\mathfrak{c}}(\bar{\mathcal{C}}) \neq 0\right\}, \tag{9.6}$$

hence by (9.5) the set in the l.h.s. of (9.6) has cardinality less than m. By applying the inductive hypothesis we finally get

$$\bar{\mathcal{C}}_u = \sum_{\mathfrak{c}} S_{\mathfrak{c}}(\bar{\mathcal{C}}_u)\mathcal{C}_{\mathfrak{c}}.$$

Putting all together we then conclude

$$\mathcal{C} = \bar{\mathcal{C}} - \mathcal{C}_{\mathfrak{c}_*} = \sum_{u=1}^{s} \bar{\mathcal{C}}_u - \mathcal{C}_{\mathfrak{c}_*} = \sum_{u=1}^{s} \sum_{\mathfrak{c}} S_{\mathfrak{c}}(\bar{\mathcal{C}}_u)\mathcal{C}_{\mathfrak{c}} - \mathcal{C}_{\mathfrak{c}_*},$$

hence $\mathcal{C}$ is a (finite) linear combination of cycles $\mathcal{C}_{\mathfrak{c}}$'s. By applying (9.2) one gets that (9.3) is satisfied. □

9.2. *Cellular homology*

Consider the graph $G_u = (V, E_u)$, for each unordered edge in E_u fix a canonical orientation and call E_o the set of canonically ordered edges (the subscript "o" stays for *ordered*, or *oriented*). In other words, E_o is any subset $E_o \subset E$ such that if $(y, z) \in E$ then either $(y, z) \in E_o$ or $(z, y) \in E_o$.

We recall the definition of the first cellular homology class $H_1(G, \mathbb{R})$ (the field $\mathbb{R}$ could be replaced by an arbitrary ring). To this aim, we introduce a proper terminology: the vertexes in V are called 0-cells and the edges in E_o are called 1-cells. For $k = 0, 1$ we define the space $C_k(\mathbb{R})$ of k-chains as the free vector space over $\mathbb{R}$ with basis given by the k-cells. Finally, we define the boundary operator

$$\partial : C_1(\mathbb{R}) \to C_0(\mathbb{R})$$

as the unique linear map such that $\partial(y, z) = z - y$ for any $(y, z) \in E_o$. The first cellular homology class $H_1(G, \mathbb{R})$ is then given by the kernel of ∂. We point out that the definition depends on the choice of the set E_o of canonically oriented edges, but any other choice of E_o would lead to a isomorphic vector space.

Since the graph has no facets of dimension 2, the family of 2-cells is empty and the space $C_2(\mathbb{R})$ of 2-chains is zero, hence the boundary operator from $C_2(\mathbb{R})$ to $C_1(\mathbb{R})$ would be the zero map. In particular, the zero 1-chain is the only exact chain, while the closed 1-chains form the kernel of the boundary operator $\partial : C_1(\mathbb{R}) \to C_0(\mathbb{R})$. Hence the above definition of $H_1(G, \mathbb{R})$ coincides indeed with the standard one, as quotient of the closed 1-chains over the exact 1-chains.

To a given cycle $\mathcal{C} = (x_1, x_2, \dots, x_k)$ in $\mathcal{V}$ we associate the homological class

$$[\mathcal{C}] := \sum_{e \in E_o} S_e(\mathcal{C}) e \tag{9.7}$$

in $H_1(G, \mathbb{R})$. Note that the above series is indeed a finite sum and that $\partial[\mathcal{C}] = 0$.

Then we have the following result:

Proposition 9.2. *The linear map* $\psi : \mathcal{V} \ni \sum_{i=1}^n a_i \mathcal{C}_i \mapsto \sum_{i=1}^n a_i [\mathcal{C}_i] \in H_1(G, \mathbb{R})$ *induces the quotient linear map*

$$\phi : \mathcal{V}_* \ni \sum_{i=1}^n a_i \mathcal{C}_i \mapsto \sum_{i=1}^n a_i [\mathcal{C}_i] \in H_1(G, \mathbb{R}),$$

which is a linear isomorphism.

Proof. To see that the map ϕ is well defined, we need to show that ψ is zero on $\mathcal{W}$ (recall that $\mathcal{V}_* = \mathcal{V}/\mathcal{W}$). To this aim, given $\sum_{i=1}^n a_i \mathcal{C}_i$ in $\mathcal{V}$ such that $\sum_{i=1}^n a_i S_e(\mathcal{C}_i) = 0$ for any $e \in E$, we have to prove that $\sum_{i=1}^n a_i [\mathcal{C}_i] = 0$. This follows easily from definition (9.7).

Let us prove that ϕ is injective. Suppose that, for some $\sum_{i=1}^n a_i \mathcal{C}_i \in \mathcal{V}_*$, it holds $\sum_{i=1}^n a_i [\mathcal{C}_i] = 0$. Since, by (9.7), $\sum_{i=1}^n a_i [\mathcal{C}_i] = \sum_{e \in E_o} \big(\sum_{i=1}^n a_i S_e(\mathcal{C}_i)\big) e$ (note that the series over $e \in E_o$ is indeed a finite sum) we conclude that $\sum_{i=1}^n a_i S_e(\mathcal{C}_i) = 0$ for any $e \in E_o$, which implies that $\sum_{i=1}^n a_i \mathcal{C}_i = 0$ in $\mathcal{V}_*$ by (9.1).

Let us prove that ϕ is surjective. To this aim fix $f = \sum_{e \in E_0} b_e e$ in $H_1(G, \mathbb{R})$ (in particular, the above series over $e \in E_o$ is a finite sum). Since $\phi(\emptyset) = 0$ we can assume $f \neq 0$. We define the flow $Q \in L^1_+(E)$ as follows: for any $e \in E_o$ with $b_e > 0$ we put $Q(e) := b_e$, while for any $e \in E_o$ with $b_e < 0$ we put $Q(\bar{e}) := -b_e$, and we set the flow Q equal to zero in all other edges. By the above definition it is simple to check that $Q(e) - Q(\bar{e}) = b_e$ for any $e \in E_o$. We now show that $\operatorname{div} Q = 0$. Indeed

$$\begin{aligned}
\operatorname{div} Q(y) &= \sum_z \big(Q(y,z) - Q(z,y)\big) \\
&= \sum_{z:(y,z) \in E_o} \big(Q(y,z) - Q(z,y)\big) + \sum_{z:(z,y) \in E_o} \big(Q(y,z) - Q(z,y)\big) \\
&= \sum_{z:(y,z) \in E_o} b_{(y,z)} - \sum_{z:(z,y) \in E_o} b_{(z,y)}.
\end{aligned}$$

On the other hand the last member equals the value of the 0-chain $-\partial f$ in y and we know that $\partial f = 0$, thus proving the zero-divergence of Q. By Lemma 4.1 in [5] and since the flow Q has finite support, we then conclude that there exist self-avoiding cycles $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_n$ and positive constants $a_1, a_2, \dots, a_n$ such that

$$Q(e) = \sum_{i=1}^n a_i \mathbb{1}(e \in \mathcal{C}_i). \tag{9.8}$$

In general we write $e \in \mathcal{C}$ if $\mathcal{C} = (x_1, x_2, \dots, x_k)$ and $e = (x_j, x_{j+1})$ for some $j \in \{1, 2, \dots, k\}$. We claim that ϕ maps $\sum_{i=1}^n a_i \mathcal{C}_i$, thought of as element of $\mathcal{V}_*$, to $f \in H_1(G; \mathbb{R})$. To this aim we need to show that

$$\sum_{i=1}^n a_i S_e(\mathcal{C}_i) = b_e \tag{9.9}$$

for any $e \in E_o$. If $b_e = 0$, then by construction $Q(e) = Q(\bar{e}) = 0$, hence by (9.8) $e, \bar{e}$ are not in the support of the $\mathcal{C}_i$'s, thus implying (9.9). If $b_e > 0$ then $Q(e) = b_e$ and $Q(\bar{e}) = 0$, hence by (9.8) $\sum_{i=1}^n a_i \mathbb{1}(e \in \mathcal{C}_i) = b_e$, while $\bar{e}$ is not in the support of the $\mathcal{C}_i$'s. This implies (9.9). If $b_e < 0$, then $Q(e) = 0$ and $Q(\bar{e}) = -b_e$, hence by (9.8) e is not in the support of the $\mathcal{C}_i$'s and $\sum_{i=1}^n a_i \mathbb{1}(\bar{e} \in \mathcal{C}_i) = -b_e$. This implies (9.9). □

9.3. LDP for the homological coefficients

Given a cycle $\mathcal{C}$ in G, its *affinity* $\mathcal{A}(\mathcal{C})$ is defined as (cf. [26])

$$\mathcal{A}(\mathcal{C}) := \sum_{j=1}^k \log \frac{r(x_j, x_{j+1})}{r(x_{j+1}, x_j)} = \sum_{j=1}^k \log \frac{\pi(x_j) r(x_j, x_{j+1})}{\pi(x_{j+1}) r(x_{j+1}, x_j)} = \sum_{j=1}^k w_\pi(x_j, x_{j+1}), \tag{9.10}$$

where $\mathcal{C} = (x_1, x_2, \dots, x_k)$ and $w_\pi : E \to \mathbb{R}$ is the function defined in (7.1). Note that we can also write

$$\mathcal{A}(\mathcal{C}) = \frac{1}{2} \sum_{e \in E} S_e(\mathcal{C}) w_\pi(e),$$

hence the above affinity induces a linear map on the cycle space $\mathcal{V}_*$.

From now on we fix a spanning tree $\mathcal{T}$ in $G_u = (V, E_u)$ and chords $\mathfrak{c}'s$. Given distinct elements $y \neq z$ in V, we call $\gamma_{y,z}$ the unique self-avoiding path $y = y_1, y_2, y_3, \dots, y_n = z$ from y to z in the tree $\mathcal{T}$.

Finally we come back to our Markov chain. To the trajectory read up to time T, $(X_t)_{0 \le t \le T}$, we associate the cycle $\mathcal{C}_T$ as follows. Let $X_0 = x_1, x_2, \dots, x_n = X_T$ be the states visited by the path $(X_t)_{0 \le t \le T}$, chronologically ordered. If $X_T = X_0$, then we set $\mathcal{C}_T := (x_1, x_2, \dots, x_{n-1})$. If $X_T \neq X_0$, then $\mathcal{C}_T := (x_1, x_2, \dots, x_n, y_1, \dots, y_{m-1})$ where $(x_n, y_1, \dots, y_m)$ is the canonical path γ_{X_T, X_0}. Roughly speaking the cycle $\mathcal{C}_T$ is obtained by gluing the trajectory $(X_t)_{0 \le t \le T}$ with the canonical path γ_{X_T, X_0} and then keeping knowledge only of the visited sites (disregarding the jump times).

Enumerating the chords as $\mathfrak{c}_k$, $k \in K$, we consider the fundamental basis $\mathcal{C}_k$, $k \in K$, where $\mathcal{C}_k := \mathcal{C}_{\mathfrak{c}_k}$. For each $k \in K$ and $T \ge 0$ we define the *empirical homological coefficient* $a_T(k)$ as the map $a_T(k) : D(\mathbb{R}_+, V) \mapsto \mathbb{R}$ characterized by the identity in $\mathcal{V}_*$

$$\mathcal{C}_T[X] = \sum_{k \in K} T a_T(k)[X] \mathcal{C}_k, \tag{9.11}$$

where $X = (X_t)_{t \in \mathbb{R}_+}$ and $\mathcal{C}_T[X]$ denotes the cycle associated to the trajectory $(X_t)_{t \in [0,T]}$. Note that we can think of a_T as a map $a_T : D(\mathbb{R}_+, V) \to L^1(K)$. We endow $L^1(K)$ with the bounded weak* topology. When K is finite this reduces to the standard L^1-topology. We write $\underline{a} = \{a(k) : k \in K\}$ for a generic element of $L^1(K)$.

Before stating our LDP for a_T we give a representation result. To this aim, for each k, let J_k be the current in $L^1_{\mathrm{a}}(E)$ satisfying $J_k(e) := S_e(\mathcal{C}_k)$ for all $e \in E$.

Lemma 9.3. *If $J \in L^1_a(E)$ has zero divergence, then $J = \sum_k J(\mathfrak{c}_k) J_k$ pointwise: $J(e) = \sum_k J(\mathfrak{c}_k) J_k(e)$ for all $e \in E$ and the series $\sum_k J(\mathfrak{c}_k) J_k(e)$ is absolutely convergent for all $e \in E$.*

Proof. Let $Q(e) := [J(e)]_+$ for any $e \in E$, where $[x]_+ := \max\{x, 0\}$. Then $Q \in L^1(E)$ and $\operatorname{div} Q = 0$. By Lemma 4.1 in [5] we can write $Q = \sum_{\mathcal{C}} \alpha_{\mathcal{C}} \mathbb{1}_{\mathcal{C}}$ with $\alpha_{\mathcal{C}} \geq 0$ and $\mathcal{C}$ varying among the self-avoiding cycles (the function $\mathbb{1}_{\mathcal{C}} : E \to \{0, 1\}$ is defined as $\mathbb{1}_{\mathcal{C}}(e) := \mathbb{1}(e \in \mathcal{C})$). Since $J(e) = Q(e) - Q(\bar{e})$ we have

$$J(e) = \sum_{\mathcal{C}} \alpha_{\mathcal{C}} \mathbb{1}(e \in \mathcal{C}) - \sum_{\mathcal{C}} \alpha_{\mathcal{C}} \mathbb{1}(\bar{e} \in \mathcal{C}),$$

and both series in the r.h.s. are convergent (and therefore absolutely convergent). Hence we can arrange the terms as we prefer and get the identities

$$J(e) = \sum_{\mathcal{C}} \alpha_{\mathcal{C}} \big(\mathbb{1}(e \in \mathcal{C}) - \mathbb{1}(\bar{e} \in \mathcal{C})\big) = \sum_{\mathcal{C}} \alpha_{\mathcal{C}} S_e(\mathcal{C}), \tag{9.12}$$

and the above series in (9.12) are absolutely convergent. By (9.3) we can write $\mathcal{C} = \sum_k S_{\mathfrak{c}_k}(\mathcal{C}) \mathcal{C}_k$, which is indeed a finite sum. In particular, $S_e(\mathcal{C})$ is given by the finite sum $\sum_k S_{\mathfrak{c}_k}(\mathcal{C}) S_e(\mathcal{C}_k)$. Coming back to (9.12) we get

$$J(e) = \sum_{\mathcal{C}} \alpha_{\mathcal{C}} \Big(\sum_k S_{\mathfrak{c}_k}(\mathcal{C}) S_e(\mathcal{C}_k)\Big). \tag{9.13}$$

Since $S_e(\mathcal{C}_k) \in \{0, -1, 1\}$ we can bound (recall that $\mathcal{C}$ is self-avoiding)

$$\begin{aligned}
\sum_{\mathcal{C}} \sum_k |\alpha_{\mathcal{C}} S_{\mathfrak{c}_k}(\mathcal{C}) S_e(\mathcal{C}_k)| &\leq \sum_{\mathcal{C}} \sum_k \alpha_{\mathcal{C}} |S_{\mathfrak{c}_k}(\mathcal{C})| \\
&= \sum_{\mathcal{C}} \sum_k \alpha_{\mathcal{C}} \big(\mathbb{1}(\mathfrak{c}_k \in \mathcal{C}) + \mathbb{1}(\bar{\mathfrak{c}}_k \in \mathcal{C})\big) \\
&= \sum_k \sum_{\mathcal{C}} \alpha_{\mathcal{C}} \mathbb{1}(\mathfrak{c}_k \in \mathcal{C}) + \sum_k \sum_{\mathcal{C}} \alpha_{\mathcal{C}} \mathbb{1}(\bar{\mathfrak{c}}_k \in \mathcal{C}) \\
&= \sum_k Q(\mathfrak{c}_k) + \sum_k Q(\bar{\mathfrak{c}}_k) \leq \|Q\|_1 < +\infty.
\end{aligned}$$

Hence the series in (9.13) is absolutely convergent, and we can rearrange its terms getting the following identities concerning absolutely convergent series (recall (9.12)):

$$J(e) = \sum_k S_e(\mathcal{C}_k) \Big(\sum_{\mathcal{C}} \alpha_{\mathcal{C}} S_{\mathfrak{c}_k}(\mathcal{C})\Big) = \sum_k S_e(\mathcal{C}_k) J(\mathfrak{c}_k) = \sum_k J_k(e) J(\mathfrak{c}_k). \quad \square$$

Due to (9.2) and (9.3) it holds

$$a_T(k) = \frac{1}{T} S_{\mathfrak{c}_k}(\mathcal{C}_T) = J_T(\mathfrak{c}_k).$$

Indeed, since $\gamma_{y,z}$ is the only self-avoiding path from y to z inside the spanning tree $\mathcal{T}$, we have $S_{\mathfrak{c}_k}\left(\gamma_{X_T, X_0}\right) = 0$. In conclusion,

$$\{a_T(k) : k \in K\} = \{J_T(\mathfrak{c}_k) : k \in K\}. \tag{9.14}$$

We have now all the tools to prove the following result (recall the definition of $J_k \in L^1_a(E)$ given before Lemma 9.3):

Theorem 9.4. *Assume the Markov chain satisfies* (A1)*,* (A2)*,* (A3) *and Condition $C(\sigma)$ with $\sigma > 0$. Then the following holds:*

(i) *As $T \to +\infty$ the sequence of probability measures $\{\mathbb{P}_x \circ a_T^{-1}\}$ on $L^1(K)$ (endowed with the bounded weak* topology) satisfies a large deviation principle with good and convex rate function $I_c : L^1(K) \to [0, +\infty]$ such that*

$$I_c(\underline{a}) = \begin{cases} \widehat{I}\left(\sum_{k\in K} a_k J_k\right) & \text{if } \sum_e \left|\sum_{k\in K} a_k J_k(e)\right| < +\infty, \\ +\infty & \text{otherwise}, \end{cases} \tag{9.15}$$

where $\widehat{I}$ is the good and convex rate function of the LDP for the empirical current obtained by contraction from Theorem 6.1.

(ii) *Suppose in addition that the function w_π introduced in (7.1) is in $C_0(E)$. If $\sum_e |\sum_{k\in K} a_k J_k(e)| = +\infty$, then it holds $I_c(\underline{a}) = I_c(-\underline{a}) = +\infty$; otherwise in $\mathbb{R} \cup \{+\infty\}$ it holds*

$$I_c(\underline{a}) = I_c(-\underline{a}) - \sum_e \Big(\sum_{k\in K} a_k J_k(e)\Big) w_\pi(e). \tag{9.16}$$

When $\sum_{e\in E} \sum_{k\in K} |a_k J_k(e)| < +\infty$, then (9.16) can be rewritten as

$$I_c(\underline{a}) = I_c(-\underline{a}) - \sum_k a_k \mathcal{A}(\mathcal{C}_k), \tag{9.17}$$

and the last series is indeed absolutely convergent.

(iii) *If the conditions of Theorem 5.2 are satisfied, then the above results remain true with $L^1(K)$ endowed with the strong L^1-topology and w_π bounded.*

Some comments on the above theorem:

Comment 1. Since $J_k(e) = \mathbb{1}(e \in \mathcal{C}_k) - \mathbb{1}(\bar{e} \in \mathcal{C}_k)$ and the cycle $\mathcal{C}_k$ is self-avoiding, we have $J_k(e) \in \{-1, 0, 1\}$ and therefore

$$\sum_{k\in K} |a_k J_k(e)| < +\infty \quad \forall e \in E \tag{9.18}$$

for any $\underline{a} \in L^1(K)$. Hence the map $E \ni e \mapsto J(e) := \sum_{k\in K} a_k J_k(e) \in \mathbb{R}$ is well defined (indeed the r.h.s. is absolutely convergent) and antisymmetric. If in addition $\sum_e |\sum_{k\in K} a_k J_k(e)| < +\infty$ then the above J belongs to $L^1(E)$, hence J is a summable current in $L^1_{\mathrm{a}}(E)$.

Comment 2. We have

$$\sum_{e\in E} \sum_{k\in K} |a_k J_k(e)| = \sum_{k\in K} |a_k| \ell(\mathcal{C}_k) \tag{9.19}$$

where $\ell(\mathcal{C}_k)$ denotes the number of edges in $\mathcal{C}_k$. Hence, the condition leading to (9.17) can be rewritten as $\sum_{k\in K} |a_k|\ell(\mathcal{C}_k) < +\infty$. It is therefore useful to know if a graph admits a fundamental basis whose cycles have uniformly bounded length. Only some partial results in this direction have been achieved in graph theory [11]. For example, working with the field $\mathbb{F}_2$ instead of $\mathbb{R}$, the following result is proved in [15]: if a locally finite transitive[2] graph has the property that the

[2] A graph G is called *transitive* if for any two vertices v, w one can exhibit a graph automorphism of G mapping v to w.

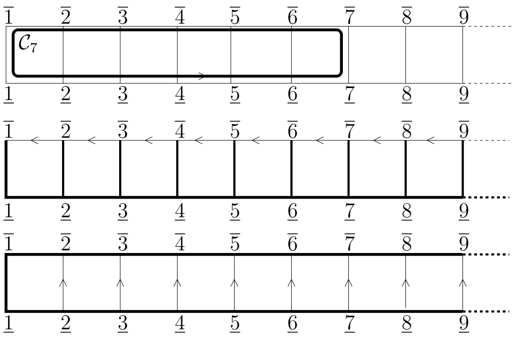

Fig. 2. *Top*. The graph G is the ladder with vertex set $\{\underline{1}, \underline{2}, \ldots\} \cup \{\overline{1}, \overline{2}, \ldots\}$, cycle $\mathcal{C}_n$ is given by $(\underline{1}, \underline{2}, \ldots, \underline{n}, \overline{n}, \overline{n-1}, \ldots, \overline{1})$. *Center*. The fundamental tree $\mathcal{T}^{(1)}$ is the bold comb. The arrows correspond to the chords. The basis cycle $\mathcal{C}_k^{(1)}$ is given by $(\overline{k+1}, \overline{k}, \underline{k}, \underline{k+1})$. *Bottom*. The fundamental tree $\mathcal{T}^{(2)}$ is in boldface. The basis cycle $\mathcal{C}_k^{(2)}$ equals $\mathcal{C}_k$.

cycle space is generated by cycles of uniformly bounded length, then the graph must be accessible (we refer to [15] for the terminology).

The lattice $\mathbb{Z}^d$ does not admit a fundamental basis whose cycles have uniformly bounded length (see the Appendix). Positive examples can be easily constructed.

Comment 3. If $\sum_{e\in E}\sum_{k\in K}|a_k J_k(e)| < +\infty$ then J has zero divergence. Indeed, given $y \in V$ we have (in the third identity we use that the series is absolutely convergent)

$$\begin{aligned}\operatorname{div} J &= \sum_z J(y,z) = \sum_z \Big(\sum_{k\in K} a_k J_k(y,z)\Big)\\ &= \sum_{k\in K} a_k \Big(\sum_z J_k(y,z)\Big) = \sum_{k\in K} a_k \cdot 0 = 0.\end{aligned}$$

Comment 4. When working with finite graphs, one can deal with an arbitrary basis of the cycle space, fixing arbitrarily once and for all the paths $\gamma_{y,z}$ from y to z and defining the homological coefficients as in (9.11) referred to the chosen basis. Then the above theorem remains true and (9.17) is always satisfied (cf. [14]).

Comment 5. For infinite graphs G, the LDP stated in Theorem 9.4 refers to the coefficients in a given basis of $H_1(G,\mathbb{R})$ and is not intrinsic to $H_1(G,\mathbb{R})$. Indeed, for suitable graphs G, by choosing different fundamental trees $\mathcal{T}_1$, $\mathcal{T}_2$ and associated fundamental cycle bases $\{\mathcal{C}_k^{(1)} : k \in \mathbb{N}_+\}$, $\{\mathcal{C}_k^{(2)} : k \in \mathbb{N}_+\}$, one can exhibit a sequence of cycles $\big(\mathcal{C}_n\big)_{n\geq 1}$ with the following property: setting $\mathcal{C}_n = \sum_{k=1}^{\infty} a_k^{(n)}\mathcal{C}_k^{(1)} = \sum_{k=1}^{\infty} b_k^{(n)}\mathcal{C}_k^{(2)}$, the n-sequence $(a_k^{(n)} : k \in \mathbb{N}_+)_{n\geq 1}$ does not converge in $L^1(\mathbb{N}_+)$ (endowed with the bounded weak* topology), while the n-sequence $(b_k^{(n)} : k \in \mathbb{N}_+)_{n\geq 1}$ does. See Fig. 2 where $(a_k^{(n)} : k \in \mathbb{N}_+)$ is the string $(1, 1, \ldots, 1, 0, 0\ldots)$ with n 1's, while $(b_k^{(n)} : k \in \mathbb{N})$ is the string $(0, 0, \ldots, 0, 1, 0, 0, \ldots)$ with a single 1 located at position n (note that $(b_k^{(n)} : k \in \mathbb{N})$ converges to the zero element of $L^1(\mathbb{N}_+)$ in the bounded weak* topology).

Proof. Item (i) as well as the first part of Item (ii) are a consequence of Lemma 9.3, identity (9.14) and the LDP for the empirical current obtained by contraction from Theorem 6.1. Let us now prove (9.17) when $\sum_{e\in E}\sum_{k\in K}|a_k J_k(e)| < +\infty$. By Item (i) we have $I_c(\underline{a}) = \widehat{I}(J)$ and $I_c(-\underline{a}) = \widehat{I}(-J)$, where $J = \sum_{k\in K} a_k J_k$ (which is indeed a summable current with zero divergence due to Problem 3). Due to (7.3) we then have

$$I_c(\underline{a}) = I_c(-\underline{a}) - \frac{1}{2}\langle J, w_\pi\rangle = I_c(-\underline{a}) - \frac{1}{2}\sum_e\Big(\sum_k a_k J_k(e)\Big)w_\pi(e).$$

Since w_π is bounded and since $\sum_{e\in E}\sum_{k\in K}|a_k J_k(e)| < +\infty$, the last series is absolutely convergent and we can rearrange it as

$$\frac{1}{2}\sum_e\Big(\sum_k a_k J_k(e)\Big)w_\pi(e) = \frac{1}{2}\sum_k a_k\sum_e\Big(J_k(e)w_\pi(e)\Big) = \sum_k a_k\mathcal{A}(\mathcal{C}_k).$$

By the same observations we also have

$$\sum_k|a_k\mathcal{A}(\mathcal{C}_k)| \le \frac{1}{2}\sum_k\Big|a_k\sum_e(J_k(e)w_\pi(e))\Big| \le \frac{\|w_\pi\|_\infty}{2}\sum_k\sum_e\big|a_k J_k(e)\big| < +\infty,$$

thus proving our thesis.

Finally, Item (iii) follows from the previous items and from Theorem 5.2. □

10. Examples

10.1. Markov chain with two states

We start by the simplest possible situation: a Markov chain with two states (a similar analysis is given in [21]). Let 0 and 1 be the two states, and denote by $r_0 = r(0,1)$ and $r_1 = r(1,0)$ the corresponding jump rates. To avoid trivialities we assume that $r_0, r_1 > 0$. The unique invariant measure π is also reversible and is given by $\pi(0) = r_1/(r_0+r_1)$, $\pi(1) = r_0/(r_0+r_1)$. Given $T > 0$ we let $q_T := Q_T(0,1) + Q_T(1,0)$ be the mean total number of jumps in the time interval $[0,T]$. We shall here derive the large deviation principle for the family of random variables $\{q_T\}_{T>0}$. We point out that the empirical current $J_T(0,1)$ is of order $O(1/T)$, hence the associated LDP is trivial.

By Theorem 3.2 and the contraction principle, the family of positive random variables $\{q_T\}_{T>0}$ satisfies a large deviation principle with rate function $f:\mathbb{R}_+\to\mathbb{R}_+$ given by

$$f(q) = \inf\big\{I(\mu,Q) : (\mu,Q)\in\mathcal{P}(V)\times L_1^+(E),\ Q(0,1)+Q(1,0) = q\big\}.$$

In view of the constraint $\operatorname{div}Q = 0$ in (3.3) we can assume $Q(0,1) = Q(1,0) = q/2$ and therefore

$$f(q) = \inf\Big\{\Phi\big(\tfrac{q}{2},\mu(0)r_0\big) + \Phi\big(\tfrac{q}{2},\mu(1)r_1\big) : \mu\in\mathcal{P}(V)\Big\}.$$

If $q = 0$ we have to minimize $\mu(0)r_0 + \mu(1)r_1$, getting therefore $f(0) = \min\{r_0, r_1\}$. If $q > 0$, writing $\mu(0) = 1/2-\gamma$ and $\mu(1) = 1/2+\gamma$, we need to minimize the function

$$\psi(\gamma) := \frac{q}{2}\log\frac{q^2}{4r_0r_1} - q + \frac{r_0+r_1}{2} + \gamma(r_1-r_0) - \frac{q}{2}\log\big(\frac{1}{4}-\gamma^2\big)$$

over $\gamma\in[-1/2, 1/2]$.

Since $\psi'(\gamma) = \left[\frac{1}{4} - \gamma^2\right]^{-1} \left[(r_0 - r_1)\gamma^2 + q\gamma + \frac{r_1 - r_0}{4}\right]$, the optimal γ is given by

$$\left[-q + \sqrt{q^2 + (r_0 - r_1)^2}\right] / 2(r_0 - r_1).$$

Hence the optimal μ is given by

$$\mu(0) = \frac{1}{2}\Big(1 + \tfrac{q}{r_0 - r_1} - \frac{\sqrt{q^2 + (r_0 - r_1)^2}}{r_0 - r_1}\Big),$$
$$\mu(1) = \frac{1}{2}\Big(1 - \tfrac{q}{r_0 - r_1} + \frac{\sqrt{q^2 + (r_0 - r_1)^2}}{r_0 - r_1}\Big),$$

understanding $\mu(0) = \mu(1) = 1/2$ when $r_0 = r_1$. In particular, we get

$$f(q) = \frac{1}{2}\Big\{q \log\Big[\frac{q}{2r_0 r_1}\big(\sqrt{q^2 + (r_0 - r_1)^2} + q\big)\Big] + r_0 + r_1 - q - \sqrt{q^2 + (r_0 - r_1)^2}\Big\}$$

and, in the special case $r_0 = r_1 = r$, $f(q) = q \log \frac{q}{r} - q + r$, which coincides with the rate function of N_T/T where N_T is a Poisson process with intensity r. Set $\overline{q} := 2r_0 r_1/(r_0 + r_1)$ and observe that, by the law of large numbers for the empirical flow, q_T converges in probability to $\overline{q}$. It is simple to check that f is a strictly convex function which achieves its minimum, as it must be the case, for $q = \overline{q}$.

10.2. *A random watch*

We consider the following random watch in which an hour consists of n minutes. At time $t = 0$ the minute hand is at 0, it stays there for an exponential time of parameter r_0 then it moves at 1,..., it stays at $n-1$ for an exponential time of parameter r_{n-1} then it moves to 0 and the hour hand advances by one,... (the exponential times are all independent). Observe that for $n > 2$ the chain just defined is not reversible while for $n = 2$ one recovers the previous 2 states Markov chain. The above random watch can be thought of also as a totally asymmetric random walk on a ring with site disorder.

Let $\mathcal{N}_T$ be the number of hours marked by such a watch in the time interval $[0, T]$. Taking the discrete torus $\mathbb{T}_n = \mathbb{Z}/n\mathbb{Z}$ as state space, note that $\mathcal{N}_T = \lfloor \sum_{i=0}^{n-1} T Q_T(i, i+1)/n \rfloor$, $\lfloor x \rfloor$ denoting the integer part of x. Hence $\mathcal{N}_T/T$ satisfies the same large deviation principle of $\sum_{i=0}^{n-1} Q_T(i, i+1)/n$. In particular, by using Theorem 3.2 and the contraction principle, we can compute the large deviation rate function f for $\mathcal{N}_T/T$. Since the only divergence-free flows are the constant flows, the rate function $f : \mathbb{R}_+ \to \mathbb{R}_+$ is given by

$$f(q) = \inf\left\{\sum_{i=0}^{n-1} \Phi(q, \mu_i r_i) : \mu_i \geq 0,\ \sum_{i=0}^{n-1} \mu_i = 1\right\}. \tag{10.1}$$

Trivially, $f(0) = \min\{r_i : 0 \leq i \leq n-1\}$. Let us assume $q > 0$. Since $\sum_{i=0}^{n-1} \Phi(q, \mu_i r_i) \geq q \sum_{i=0}^{n-1} \log \frac{1}{\mu_i} - C$ for a suitable constant C independent from $\{\mu_i\}$, we conclude that the above infimum is indeed achieved inside the region $\{\mu_i > 0\ \forall i\}$. Introducing the Lagrangian multiplier λ, we first look for the extremal points of

$$\psi\left(\{\mu_i\}, \lambda\right) = \sum_{i=0}^{n-1} \Phi(q, \mu_i r_i) + \lambda\left(\sum_{i=0}^{n-1} \mu_i - 1\right).$$

These are characterized by the system

$$\begin{cases} -\dfrac{q}{\mu_i} + r_i + \lambda = 0, \\ \displaystyle\sum_{i=0}^{n-1} \mu_i = 1. \end{cases} \tag{10.2}$$

We restrict to the region $\{\mu_i > 0\ \forall i\}$ as it must be. From the first identity we get that $\lambda > -r_{\min}$ where $r_{\min} := \min_i r_i$. Let $R : (-r_{\min}, +\infty) \to (0, +\infty)$ be the strictly increasing function defined by

$$\frac{1}{R(\lambda)} = \sum_{i=0}^{n-1} \frac{1}{r_i + \lambda}.$$

We denote by $R^{-1} : (0, +\infty) \to (-r_{\min}, +\infty)$ the corresponding inverse function. Then the unique solution of (10.2) is given by $\lambda = R^{-1}(q)$ and $\mu_i = q/(R^{-1}(q) + r_i)$. This gives also the minimizer of (10.1). In particular, the large deviation rate function $f : \mathbb{R}_+ \to [0, +\infty)$ associated to $\mathcal{N}_T/T$ is given by

$$f(q) = \sum_{i=0}^{n-1} q \log\big(1 + \tfrac{R^{-1}(q)}{r_i}\big) - R^{-1}(q)$$

where we understand $f(0) = r_{\min}$.

Note that the invariant measure π_i is given by $\pi_i = r_i^{-1} / \sum_{k=0}^{n-1} r_k^{-1}$. Hence, $\mathcal{N}_T/T$ converges in probability to $n^{-1} \sum_{i=0}^{n-1} \pi_i r_i = R(0)$. Indeed, we have $f(R(0)) = 0$ as it must be.

Finally, we point out that $J_T(i, i+1) = Q_T(i, i+1)$, hence the large deviations for the current and for the flow coincide.

10.3. One particle on a ring

Consider a homogeneous simple random walk on the discrete one dimensional torus with N sites $\mathbb{T}_N := \mathbb{Z}/N\mathbb{Z}$. The generator of the process is

$$L_N f(x) = \lambda p\big[f(x+1) - f(x)\big] + \lambda(1-p)\big[f(x-1) - f(x)\big], \tag{10.3}$$

where $x \in \mathbb{T}_N$, λ is a positive parameter and $p \in [0, 1]$. We are interested in the rate function for the empirical current $J_T(x, x+1) = Q_T(x, x+1) - Q_T(x+1, x)$. By symmetry the rate function does not depend on x (we refer to [23] for related results).

The rate function can be computed directly since it coincides with the rate function of X_T/N where X_T is a simple random walk on $\mathbb{Z}$ having generator (10.3). Indeed, if for example the random walk starts at x, $\lfloor X_T/N \rfloor$ corresponds to the number of cycles made by the walker, with the rule that a clockwise cycle has weight 1 and a unclockwise cycle has weight -1. In particular, $T J_T(x, x+1)$ differs from $\lfloor X_T/N \rfloor$ by at most one, hence $|J_T(x, x+1) - X_T/NT| \le 2/T$. Note that for $p = \pm 1$, we have $X_T = \pm N_T$, $(N_t)_{t \in \mathbb{R}_+}$ being a Poisson process of parameter λ. To simplify the treatment below, we restrict to $p \in (0, 1)$ excluding the trivial cases $p = \pm 1$.

The rate function of X_T/N can be easily computed by means of Gärtner–Ellis Theorem using the representation

$$X_T = \sum_{i=1}^{N_T} Y_i,$$

where $(N)_{t\in\mathbb{R}_+}$ is a Poisson process of parameter λ and Y_i are independent i.i.d. random variables taking values $1, -1$ with probability $p, 1-p$, respectively. We have that the corresponding rate function W_N is obtained as

$$W_N(j) = \sup_{\alpha\in\mathbb{R}} \{j\alpha - \Lambda_N(\alpha)\}, \tag{10.4}$$

where

$$\begin{aligned}\Lambda_N(\alpha) &= \lim_{T\to+\infty} \frac{1}{T}\log\mathbb{E}\left(e^{\frac{\alpha X_T}{N}}\right)\\ &= \lim_{T\to+\infty}\frac{1}{T}\log\left(\sum_{k=0}^{+\infty} e^{-\lambda T}\frac{(\lambda T)^k}{k!}\mathbb{E}\left(e^{\frac{\alpha}{N}(Y_1+\cdots+Y_k)}\right)\right)\\ &= \lim_{T\to+\infty}\frac{1}{T}\log\left(\sum_{k=0}^{+\infty} e^{-\lambda T}\frac{(\lambda T)^k}{k!}\left(pe^{\frac{\alpha}{N}}+(1-p)e^{-\frac{\alpha}{N}}\right)^k\right)\\ &= \lambda\, p\, e^{\frac{\alpha}{N}} + \lambda(1-p)e^{-\frac{\alpha}{N}} - \lambda. \end{aligned}\tag{10.5}$$

Putting (10.5) into (10.4), one gets that the supremum in (10.4) is attained at $\alpha = N\log\left(Nj/2p\lambda + (1/2p\lambda)\sqrt{(Nj)^2+4p(1-p)\lambda^2}\right)$, hence

$$\begin{aligned}W_N(j) &= Nj\log\left(\frac{Nj}{2p\lambda}+\frac{1}{2p\lambda}\sqrt{(Nj)^2+4p(1-p)\lambda^2}\right)\\ &\quad -\sqrt{(Nj)^2+4p(1-p)\lambda^2}+\lambda.\end{aligned}\tag{10.6}$$

Note that, given $F\in\mathbb{R}$, choosing $p=\frac{1}{2}+\frac{F}{2N}$ and $1-p=\frac{1}{2}-\frac{F}{2N}$ (for N large enough) and $\lambda=\gamma N^2$, it holds

$$\lim_{N\to+\infty} W_N(j) = \frac{(j-\gamma F)^2}{2\gamma}.$$

The above asymptotics is consistent with formula (58) in [22] of the large deviation rate function for the current of a diffusion on the circle.

The same result, i.e. the LD rate functional for $J_T(x, x+1)$, can be obtained by a purely variational approach. We write J for the unique zero divergence current such that $J(x, x+1) = j$. By Theorem 6.1 and the contraction principle, we get

$$W_N(j) = \inf\left\{\widetilde{I}(\mu, J) : \mu\in\mathcal{P}(V)\right\},$$

where $\widetilde{I}(\mu, J)$ has been defined in (6.5). Since $\widetilde{I}(\cdot, J)$ is l.s.c. on the compact space $\mathcal{P}(V)$, the above infimum is obtained at some minimizer. We call Γ the set of minimizers $\mu\in\mathcal{P}(V)$ and observe that Γ is convex since $\widetilde{I}$ is convex. As $\widetilde{I}(\cdot, J)$ is left invariant by the transformation $\mu\to\mathcal{T}\mu$ with $\mathcal{T}\mu = \{\mu_{y+1}\}_{y\in\mathbb{T}_N}$ if $\mu = \{\mu_y\}_{y\in\mathbb{T}_N}$, also Γ is $\mathcal{T}$-invariant. Fix $\mu\in\Gamma$. Then, $\mu, \mathcal{T}\mu, \mathcal{T}^2\mu, \dots, \mathcal{T}^{N-1}\mu$ all belong to Γ. By convexity of Γ, the uniform measure $\mu_* = \frac{1}{N}\sum_{j=0}^{N-1}\mathcal{T}^j\mu$ is in Γ. Hence, $W_N(j) = \widetilde{I}(\mu_*, J)$ and from (6.5) one recovers (10.6).

10.4. Birth and death chains

Consider the birth and death Markov chain on $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ with rates $r(k, k+1) = b_k > 0$ for $k \geq 0$ and $r(k, k-1) = d_k > 0$ for $k \geq 1$. This chain has been treated in details in [5]. Here we restrict to investigate when the joint LDP for the empirical measure and flow holds with the L^1-strong topology instead of the bounded weak* topology.

As proved in [5, Section 9], if $\lim_{k\to\infty} d_k = +\infty$ and $\overline{\lim}_{k\to\infty} b_k/d_k < 1$, then Condition $C(\sigma)$ is satisfied for some $\sigma > 0$ (as well the basic assumptions (A1),…,(A4)). Then the following holds

Proposition 10.1. *Suppose that* $\lim_{k\to\infty} d_k = +\infty$ *and* $\overline{\lim}_{k\to\infty} b_k/d_k < 1$.

(i) *If* $\lim_{k\to\infty} b_k/d_k = 0$, *then the joint LDP for* (μ_T, Q_T) *holds with* $L^1_+(E)$ *endowed with the strong topology;*

(ii) *If* $\underline{\lim}_{k\to\infty} b_k/d_k > 0$, *then the joint LDP for* (μ_T, Q_T) *does not hold with* $L^1_+(E)$ *endowed with the strong topology.*

Proof. We first derive (i) by applying Theorem 5.2 to which we refer for the notation. We define $\widehat{E} := \{(k, k+1) : k \in \mathbb{Z}_+\}$. Then $H(k) = b_k/(b_k + d_k)$ so that, by assumption, $\lim_{k\to\infty} H(k) = 0$. Hence, Items (i) and (ii) of Theorem 5.2 are satisfied.

Given $a > 0$ and a state $x \in \mathbb{Z}_+$, we choose $k_* \geq x$ such that $H(k) < a$ for any $k \geq k_*$ and define $W = W(x, a) := \{(k, k+1),\ (k+1, k),\ k \geq k_*\}$. In particular, Items (iii.1) and (iii.2) in Theorem 5.2 are satisfied.

It remains to check Item (iii.3). For any path exiting from x, given $k \geq k_*$ we get that the number of times the path uses the edge $(k, k+1)$ is at least the number of times the path uses the edge $(k+1, k)$ (more precisely, we have equality when the path ends inside $[0, k] \cap \mathbb{Z}_+$ while we have a difference of one unit if the path ends outside $[0, k]$). In conclusion (5.2) is valid with $\gamma = 1/2$.

To prove Item (ii) we generalize the argument used at the end of Section 9 in [5]. We restrict to n large enough that $1/d_n + 1/d_{n+1} < 1$. In this case we define

$$\begin{aligned}
\gamma_n &:= 1 - \frac{1}{d_n} - \frac{1}{d_{n+1}},\\
\mu^n &:= \gamma_n \pi + \frac{\delta_n}{d_n} + \frac{\delta_{n+1}}{d_{n+1}},\\
Q^n &:= \gamma_n Q^\pi + \delta_{(n,n+1)} + \delta_{(n+1,n)}.
\end{aligned}$$

Note that Q^n is divergence-free. For all edges (y, z) different from $(n, n-1)$, $(n, n+1)$, $(n+1, n)$, $(n+1, n+2)$ it holds $\Phi\big(Q^n(y, z), \mu^n(y) r(y, z)\big) = 0$ since $Q^n(y, z) = \mu^n(y) r(y, z)$.

On the other hand

$$\begin{aligned}
\Phi\big(Q^n(n, n-1), \mu^n(n) r(n, n-1)\big) &= \Phi\big(q_n^{(1)}, p_n^{(1)}\big),\\
\Phi\big(Q^n(n, n+1), \mu^n(n) r(n, n+1)\big) &= \Phi\big(q_n^{(2)}, p_n^{(2)}\big),\\
\Phi\big(Q^n(n+1, n), \mu^n(n+1) r(n+1, n)\big) &= \Phi\big(q_n^{(3)}, p_n^{(3)}\big),\\
\Phi\big(Q^n(n+1, n+2), \mu^n(n+1) r(n+1, n+2)\big) &= \Phi\big(q_n^{(4)}, p_n^{(4)}\big),
\end{aligned}$$

where

$$\begin{aligned}
q_n^{(1)} &:= \gamma_n Q^\pi(n, n-1), & p_n^{(1)} &:= \gamma_n Q^\pi(n, n-1) + 1,\\
q_n^{(2)} &:= \gamma_n Q^\pi(n, n+1) + 1, & p_n^{(2)} &:= \gamma_n Q^\pi(n, n+1) + \frac{b_n}{d_n},
\end{aligned}$$

$$q_n^{(3)} := \gamma_n Q^\pi(n+1,n)+1, \qquad p_n^{(3)} := \gamma_n Q^\pi(n+1,n)+1,$$
$$q_n^{(4)} := \gamma_n Q^\pi(n+1,n+2), \qquad p_n^{(4)} := \gamma_n Q^\pi(n+1,n+2)+\frac{b_{n+1}}{d_{n+1}}.$$

Trivially, $\Phi\big(q_n^{(3)}, p_n^{(3)}\big) = 0$. For $p \geq q$ we have $0 \leq \Phi(q, p) \leq p - q$; hence $\Phi\big(q_n^{(i)}, p_n^{(i)}\big)$ is uniformly bounded for $i = 1, 4$. Since $\underline{\lim}_{k\to\infty} b_k/d_k \in (0, 1)$, we can extract a subsequence $\{n_k\}_{k\geq 1}$ such that $0 < c \leq b_{n_k}/d_{n_k} \leq c'$ for some fixed $c, c' > 0$ and for all $k \geq 1$. As Q^π is summable then $\gamma_n Q^\pi(n, n+1)$ is uniformly bounded. We conclude that $\sup_{k\geq 1} \Phi\big(q_{n_k}^{(2)}, p_{n_k}^{(2)}\big) < +\infty$.

We have thus shown that $\overline{\lim}_{k\to\infty} I(\mu^{n_k}, Q^{n_k}) < +\infty$. We cannot therefore have a LDP with $L^1_+(E)$ endowed with the strong topology since the level sets of I would be compact while the sequence $\big\{(\mu^{n_k}, Q^{n_k})\big\}_{k\geq 1}$ is not relatively compact in $L^1_+(E)$ with the strong topology. □

Remark 10.2. Since the only current associated to the birth–death chain with vanishing divergence is the zero current, the LDP for the empirical current becomes trivial.

10.5. Random walks with confining potential and external force

We now apply some of our previous considerations to the nearest neighbor random walk on $\mathbb{Z}^d$ with jump rates

$$r(y,z) = \exp\Big\{-\frac{1}{2}\big[U(z)-U(y)\big]+\frac{1}{2}F(y,z)\Big\}, \quad (y,z)\in E, \tag{10.7}$$

where $E := \big\{(y, z) \in \mathbb{Z}^d \times \mathbb{Z}^d, |x - y| = 1\big\}$, $U : \mathbb{Z}^d \to \mathbb{R}$ is a function satisfying $\sum_{y\in\mathbb{Z}^d} \exp\{-U(y)\} < +\infty$ (in particular U has compact level sets) and $F \in L^\infty(E)$. It is convenient to set

$$r_0(y,z) = \exp\Big\{-\frac{1}{2}\big[U(z)-U(y)\big]\Big\}, \quad (y,z)\in E. \tag{10.8}$$

Note that when $r(\cdot, \cdot) = r_0(\cdot, \cdot)$, the random walk is reversible with respect to the probability $\pi = \exp\{-U\}$, where we assume that U has been chosen so that π is properly normalized. As usual, we denote by r the holding time parameters, i.e. $r(y) = \sum_{z\sim y} r(y, z)$ where the summation is carried out over the nearest neighbors of y.

If one regards the random walk with rates (10.7) as a model for the position of a charged particle in the confining potential U, the function F is naturally interpreted as the external field.

We start discussing explosion, i.e. Assumption (A.2). A sufficient condition for non explosion is given by Theorem 4.6 in [28]: explosion does not occur if there exist a constant $\gamma \geq 0$ and a nonnegative function G such that $G(x_n) \to +\infty$ when $r(x_n) \to +\infty$ and such that (recall (3.1))

$$LG(y) \leq \gamma G(y), \quad \forall y \in \mathbb{Z}^d. \tag{10.9}$$

Consider the function $G(y) = e^{\frac{U(y)}{2}}$. This is nonnegative and has compact level sets. We have

$$\sum_z r(y,z)\Big(G(z)-G(y)\Big) \leq \sum_z r(y,z)G(z)$$
$$= G(y)\sum_z e^{\frac{F(y,z)}{2}} \leq 2de^{\frac{\|F\|_\infty}{2}}G(y).$$

We therefore conclude that explosion never occurs.

To continue our investigation of the other assumptions, we consider the radial and the transversal variation of the potential. More precisely, when $U \in C^1(\mathbb{R}^d)$ we consider the orthogonal decomposition

$$\nabla U(y) = \langle \nabla U(y), \widehat{y} \rangle \, \widehat{y} + W(y), \qquad y \in \mathbb{R}^d \setminus \{0\} \tag{10.10}$$

with $\widehat{y} := y/|y|$ and $\langle y, W(y) \rangle = 0$. Above $\langle \cdot, \cdot \rangle$ is the inner product in $\mathbb{R}^d$.

We say that the potential $U \in C^1(\mathbb{R}^d)$ has *diverging radial variation which dominates the transversal variation* if

$$\lim_{|y| \to \infty} \langle \nabla U(y), \widehat{y} \rangle = +\infty, \tag{10.11}$$

and

$$|W(y)| \le \frac{\alpha}{\sqrt{d}} \langle \nabla U(y), \widehat{y} \rangle + C, \tag{10.12}$$

for some $\alpha \in [0, 1)$ and some $C \ge 0$. Note that if W in (10.10) is bounded, then (10.12) is trivially satisfied with $\alpha = 0$. Moreover, note that (10.11) implies that $\lim_{|y| \to \infty} U(y)/|y| = +\infty$.

We give a criterion assuring Condition $C(\sigma)$.

Lemma 10.3 (Condition $C(\sigma)$). *If* $\lim_{|y| \to \infty} r_0(y) = +\infty$, *then Condition* $C(\sigma)$ *holds for some* $\sigma > 0$. *In particular, if* $U \in C^1(\mathbb{R}^d)$ *has* diverging radial variation which dominates the transversal variation, *then Condition* $C(\sigma)$ *holds for some* $\sigma > 0$.

Proof. We first prove the first part. As u_n we pick the constant sequence $u = \exp\{U/2\}$. Items (i)–(iv) in Condition $C(\sigma)$ then hold trivially. Moreover,

$$\begin{aligned} v(y) &= -\frac{Lu}{u}(y) = \sum_{z : z \sim y} r(y, z) - \sum_{z : z \sim y} \exp\left\{\tfrac{1}{2} F(y, z)\right\} \\ &\ge r(y) - 2d \, \exp\left\{\tfrac{1}{2} \|F\|_\infty\right\} \ge r_0(y) \exp\left\{-\tfrac{1}{2} \|F\|_\infty\right\} - 2d \, \exp\left\{\tfrac{1}{2} \|F\|_\infty\right\} \end{aligned}$$

which imply Items (v) and (vi).

Let now U be as in the second part of the lemma. Fix $y \in \mathbb{Z}^d \setminus \{0\}$. There must exist a unit vector $e \in \mathbb{Z}^d$ such that $\langle y, e \rangle \ge |y|/\sqrt{d}$. Set $z = y - e$. Then, for some $\xi = y - se$ and $s \in [0, 1]$, we can write

$$\begin{aligned} U(y) - U(z) &= \langle \nabla U(\xi), e \rangle = \langle \nabla U(\xi), \widehat{\xi} \rangle \langle \widehat{\xi}, e \rangle + \langle W(\xi), e \rangle \\ &\ge \langle \nabla U(\xi), \widehat{\xi} \rangle \left[\langle \widehat{\xi}, e \rangle - \frac{\alpha}{\sqrt{d}} \right] - C, \end{aligned}$$

where in the last bound we used (10.12). Since $\langle \xi, e \rangle = \langle y, e \rangle - s \ge |y|/\sqrt{d} - 1$ while $|\xi| \le |y| + 1$, we conclude that

$$U(y) - U(z) \ge \frac{\langle \nabla U(\xi), \widehat{\xi} \rangle}{\sqrt{d}} \left[\frac{|y| - \sqrt{d}}{|y| + 1} - \alpha \right] - C. \tag{10.13}$$

The above inequality gives a lower bound for $r_0(y, z)$, and therefore for $r_0(y)$, which implies that $\lim_{|y| \to \infty} r_0(y) = +\infty$ under assumption (10.11). □

We now give a criterion assuring that the joint LDP of Theorem 3.2 holds with $L^1_+(E)$ endowed with the strong L^1-topology.

Lemma 10.4 (*LDP in L^1-strong topology*). *Suppose that $U \in C^1(\mathbb{R}^d)$ has* diverging radial variation which dominates the transversal variation. *Consider one of the two following cases:*

Case 1*: W is bounded (which automatically implies (10.12));*

Case 2*: (10.12) holds for some $\alpha \in [0, 1/2)$ and*

$$\lim_{\substack{|y|,|z|\to\infty\\ |y-z|\le 1}} \frac{\langle \nabla U(y), \widehat{y}\,\rangle}{\langle \nabla U(z), \widehat{z}\,\rangle} = 1. \tag{10.14}$$

Then, both in Case 1 *and in Case* 2*, Theorem* 3.2 *holds with $L^1_+(E)$ endowed with the strong L^1-topology.*

Proof. We apply Theorem 5.2 with

$$\widehat{E} := \{(y, y+e) \in \mathbb{Z}^d \times \mathbb{Z}^d : |e| = 1,\ \langle y, e\rangle \ge 0\}.$$

The validity of Item (i) of Theorem 5.2 is trivial. Let us check Item (ii) of Theorem 5.2. We restrict to Case 2 (Case 1 follows the main lines and is simpler, we give some comments below). To this aim fix $y \in \mathbb{Z}^d \backslash \{0\}$. Take $z \in \mathbb{Z}^d$ with $z = y + e$, $|e| = 1$ and $\langle y, e\rangle \ge 0$. Then, for some $\xi = y + se$ and $s \in [0, 1]$, we can write

$$U(z) - U(y) = \langle \nabla U(\xi), e\rangle = \langle \nabla U(\xi), \widehat{\xi}\,\rangle\langle \widehat{\xi}, e\rangle + \langle W(\xi), e\rangle.$$

Since $\langle \xi, e\rangle = \langle y, e\rangle + s \ge 0$, for $|y|$ large we can bound

$$U(z) - U(y) \ge -|W(\xi)| \ge -\frac{\alpha}{\sqrt{d}}\langle \nabla U(\xi), \widehat{\xi}\,\rangle - C.$$

This implies for $|y|$ large that

$$\sum_{z:(y,z)\in\widehat{E}} r(y, z) \le 2d \exp\left\{\frac{\alpha\gamma_+(y)}{2\sqrt{d}} + \frac{\|F\|_\infty + C}{2}\right\}, \tag{10.15}$$

where

$$\gamma_+(y) := \sup\left\{\langle \nabla U(\xi), \widehat{\xi}\,\rangle : \xi \in \mathbb{R}^d,\ |\xi - y| \le 1\right\}.$$

In Case 1 (10.15) remains valid with $\frac{\alpha\gamma_+(y)}{2\sqrt{d}}$ replaced by $\sup_{i:1\le i\le d}\|W_i\|_\infty/2$, where $W = (W_1, \dots, W_d)$.

Take e' a unit vector such that $\langle y, e'\rangle \ge |y|/\sqrt{d}$ and set $z' = z - e'$. Being in the same setting of (10.13), we conclude that

$$r(y, z') \ge \exp\left\{\frac{\gamma_-(y)}{2\sqrt{d}}\left[\frac{|y| - \sqrt{d}}{|y| + 1} - \alpha\right] - \frac{C}{2} - \frac{1}{2}\|F\|_\infty\right\}, \tag{10.16}$$

where

$$\gamma_-(y) := \inf\left\{\langle \nabla U(\xi), \widehat{\xi}\,\rangle : \xi \in \mathbb{R}^d,\ |\xi - y| \le 1\right\}.$$

By using (10.15) and (10.16) we get

$$H(y) \le \frac{\sum_{z:(y,z)\in\widehat{E}} r(y, z)}{r(y, z')} \le C' \exp\left\{\frac{\gamma_-(y)}{2\sqrt{d}}\left(\alpha\frac{\gamma_+(y)}{\gamma_-(y)} + \alpha - \frac{|y| - \sqrt{d}}{|y| + 1}\right)\right\}. \tag{10.17}$$

Since the map $\xi \to \langle \nabla U(\xi), \widehat{\xi}\,\rangle$ is continuous, we can write $\gamma_+(y) = \langle \nabla U(\xi_0), \widehat{\xi_0}\,\rangle$ and $\gamma_-(y) = \langle \nabla U(\xi_1), \widehat{\xi_1}\,\rangle$ for suitable ξ_0, ξ_1 satisfying $|\xi_0 - y|, |\xi_1 - y| \le 1$. Writing

$$\frac{\gamma_+(y)}{\gamma_-(y)} = \frac{\langle \nabla U(\xi_0), \widehat{\xi_0}\,\rangle}{\langle \nabla U(y), \widehat{y}\,\rangle} \frac{\langle \nabla U(y), \widehat{y}\,\rangle}{\langle \nabla U(\xi_1), \widehat{\xi_1}\,\rangle},$$

by (10.14) we deduce that $\gamma_+(y)/\gamma_-(y) = 1 + o(1)$ as $|y| \to +\infty$. In particular, we can rewrite (10.17) as

$$H(y) \le C' \exp\left\{ \frac{\gamma_-(y)}{2\sqrt{d}} \big(2\alpha - 1 + o(1)\big) \right\}.$$

Using that $\gamma_-(y) \to +\infty$ as $|y| \to +\infty$ and that $\alpha < 1/2$ (we restrict to Case 2), we get Item (ii) of Theorem 5.2, i.e. that the function H defined in (5.1) vanishes at infinity.

Let us finally check Item (iii) of Theorem 5.2. To this aim, given a positive integer r, we introduce the diamond $B(r) := \{y \in \mathbb{Z}^d : |y|_1 \le r\}$. Given $x \in \mathbb{Z}^d$ and $a > 0$, we take r large enough that $x \in B(r)$ and $\{H \ge a\} \subset B(r-1)$ (recall that H vanishes at infinity). Finally we define $W = W(x, a)$ as the family of oriented edges in $\mathbb{Z}^d$ not inside $B(r)$:

$$W := \{(y, z) \in E : y \notin B(r) \text{or } z \notin B(r)\}.$$

Trivially W satisfies Items (iii.1) and (iii.2) in Theorem 5.2. We claim that also Item (iii.3) holds: given any path $x_1 = x, x_2, x_3 \dots x_n$ of nearest-neighbor points in $\mathbb{Z}^d$ starting at x, the number of its edges in $W \cap \widehat{E}$ is at least $1/2$ of the total number of its edges in W. To prove the above claim it is enough to observe that, considering the pieces of the path in $\{y \in \mathbb{Z}^d : |y|_1 \ge r\}$, we can restrict to a path $x_1, x_2, x_3 \dots x_n$ with $|x_1|_1 = r$ and with $|x_i|_1 \ge r$ for all $i = 2, \dots, n$. To prove the thesis for this path, we observe that $|x_{i+1}|_1 = |x_i|_1 + 1$ if $x_{i+1} - x_i \in \widehat{E}$ while $|x_{i+1}|_1 = |x_i|_1 - 1$ if $x_{i+1} - x_i \notin \widehat{E}$. Therefore,

$$\begin{aligned} &\sharp\{i : 1 \le i < n,\ x_{i+1} - x_i \in \widehat{E}\} - \sharp\{i : 1 \le i < n,\ x_{i+1} - x_i \notin \widehat{E}\} \\ &\quad = |x_n|_1 - |x_1|_1 = |x_n|_1 - r. \end{aligned}$$

Since by assumption $|x_n|_1 \ge r$ we get the thesis. $\square$

We next discuss some choices of the field F allowing to apply Theorem 8.1 and to deduce the large deviation principle for the Gallavotti–Cohen functional. These hypotheses will be in the same spirit of those introduced in [4] for continuous diffusions. Observing that in this example it holds $E = E_s$, we restrict to the physically relevant case in which F is antisymmetric, i.e. $F(y, z) = -F(z, y)$, $(y, z) \in E$. We then require that the chain with rates r has the same invariant measure $\pi = \exp\{-U\}$ as the one with rates r_0, that is

$$\sum_{z : z \sim y} \exp\left\{-\tfrac{1}{2}[U(z) + U(y)]\right\} \sinh\left(\frac{1}{2} F(y, z)\right) = 0, \qquad \forall y \in V. \tag{10.18}$$

We stress that the knowledge of π is necessary to know the function w_π in (7.1), we consider here models where the external force field does not change the invariant distribution.

For simplicity we restrict to $d = 2$. Functions U and F satisfying (10.18) can be easily constructed. For instance one can take U "radial", i.e. $U(y) = \widetilde{U}(|y|_1)$ for some $\widetilde{U} : \mathbb{Z}_+ \to \mathbb{R}$. Then the discrete vector field F has to be fixed as in Fig. 3. In that figure we represent the level curves of U with black lines and use arrows of different colors to represent the force field. To each color we arbitrarily associate a real number varying in a fixed interval $[-A, A]$ representing the value of the discrete vector field. Consider an oriented edge (y, z). If in Fig. 3 there is a

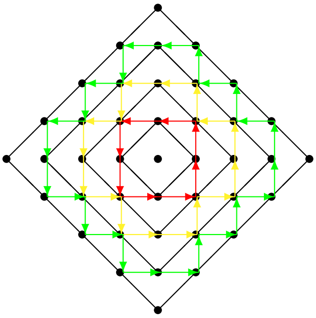

Fig. 3. The vector field F when $U = \widetilde{U}(|x|_1)$. (The reader is referred to the web version of this article for colored figures.)

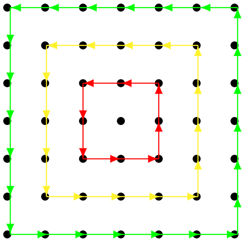

Fig. 4. The vector field F when $U = \widetilde{U}(|x|_\infty)$. (The reader is referred to the web version of this article for colored figures.)

colored arrow from y to z then $F(y,z)$ assumes the value corresponding to that color, while if there is a colored arrow from z to y then $F(y,z)$ assumes the value corresponding to that color with a minus sign. If there is no arrow associated either to (y,z) or to (z,y) then $F(y,z) = 0$. Note that by construction $\|F\|_\infty$ is bounded and (10.18) is satisfied.

If instead we consider U of the form $U(y) = \widetilde{U}(|x|_\infty)$ for some $\widetilde{U} : \mathbb{Z}_+ \to \mathbb{R}$ then to have (10.18) we need to fix the discrete vector field F as in Fig. 4, following the same construction as above. In both cases the discrete vector field F is associated to “rotations” along the level curves of U.

The Gallavotti–Cohen functional (8.2) then becomes

$$W_T = \frac{1}{2} \sum_{(y,z)\in E} J_T(y,z)F(y,z) = \sum_{y\in\mathbb{Z}^d} \sum_{i=1}^{d} J_T(y, y+e_i)F(y, y+e_i)$$

where we used the antisymmetry of J_T and F. In particular, W_T is naturally interpreted as the empirical power dissipated by F. The large deviation principle for the family $\{W_T\}$ then follows from Theorem 8.1. In particular, if $F \in C_0(E)$ we only need to require Condition $C(\sigma)$ for some $\sigma > 0$ and this can be checked using the criterion given in Lemma 10.3. If $F \in L^\infty(E)$ we need in addition to verify that the joint LDP for the empirical measure and flow holds with the L^1-topology instead of the bounded weak* topology for $L^1_+(E)$. This can be done by applying Theorem 5.2, or the criterion (as well as some variations) given in Lemma 10.4.

Acknowledgments

We thank R. Diestel for useful discussions. We thank the anonymous referee for his/her helpful comments and corrections.

Appendix. Geometric properties of spanning trees of $\mathbb{Z}^d$

We consider here the lattice $\mathbb{Z}^d$, $d \geq 2$. Trivially, the cycle space admits a basis given by cycles of uniformly bounded length: take the cycles $(x, x + e_i, x + e_i + e_j, x + e_j)$ where x varies in $\mathbb{Z}^d$, $1 \leq i < j \leq d$, e_i and e_j vary among the vectors in the canonical basis of $\mathbb{Z}^d$. Due to Problem 2 after Theorem 9.4 it is natural to ask if the lattice $\mathbb{Z}^d$ admits a fundamental basis given by cycles of uniformly bounded length. The answer is negative due to the following fact:

Proposition A.1. *Consider a countable connected unoriented graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and fix a spanning tree $\mathcal{T}$. If the fundamental cycle basis associated to $\mathcal{T}$ has cycles with at most $\ell + 1$ vertices, then the following property holds:*

Given $a \neq b \in \mathcal{V}$ fix a path $\gamma = (x_0, x_1, \dots, x_M)$ from $x_0 = a$ to $x_M = b$. Let $\gamma_{a,b} = (z_0, z_1, \dots, z_R)$ be the unique self-avoiding path inside the tree $\mathcal{T}$ from $z_0 = a$ to $z_R = b$. Then for any $i : 0 \leq i \leq R$ there exists $j : 0 \leq j \leq M$ with $d(z_i, x_j) \leq \ell$, $d(\cdot, \cdot)$ being the graph distance.

Since the property in the above proposition is trivially not satisfied by the lattice $\mathbb{Z}^d$, $d \geq 2$, we get that $\mathbb{Z}^d$ has no fundamental cycle basis with uniformly bounded length.

Proof. Consider the path $\gamma = (x_0, x_1, x_2, \dots, x_M)$. For each $k = 0, 1, \dots, M - 1$ either the edge (x_k, x_{k+1}) belongs to the tree $\mathcal{T}$, or it is a chord and therefore the vertices x_k, x_{k+1} have graph distance bounded by ℓ inside $\mathcal{T}$. We modify γ as follows. If the edge (x_k, x_{k+1}) belongs to the tree $\mathcal{T}$, then keep the pair x_k, x_{k+1} unchanged, otherwise replace the pair x_k, x_{k+1} by the string $x_k, a_1, a_2, \dots, a_r, x_{k+1}$ given by the unique self-avoiding path inside $\mathcal{T}$ from x_k to x_{k+1}. We call $\gamma^{(1)}$ the resulting new path. Writing $\gamma^{(1)} = (y_0, y_1, \dots, y_S)$, we get that $y_0 = a$, $y_S = b$, $\gamma^{(1)}$ lies inside the tree and that

$$\forall i : 0 \leq i \leq S \quad \exists j : 0 \leq j \leq M \text{ such that } d(y_i, x_j) \leq \ell. \tag{A.1}$$

The path $\gamma^{(1)}$ could have self-intersections, anyway thought of as an unoriented graph it is a connected subgraph of $\mathcal{T}$, hence it contains a self-avoiding path from a to b, which (by definition of tree) must be $\gamma_{a,b}$. In particular, the vertices of $\gamma_{a,b}$ are of the form y_i and therefore satisfy (A.1). □

References

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

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

[3] M. Baiesi, C. Maes, K. Netŏcný, Computation of current cumulants for small nonequilibrium systems, J. Stat. Phys. 135 (2009) 57–75.

[4] L. Bertini, G. Di Gesù, Small noise asymptotic of the Gallavotti–Cohen functional in diffusion processes, ALEA (2015) in press.

[5] L. Bertini, A. Faggionato, D. Gabrielli, Large deviations for the empirical flow for continuous time Markov chains, Ann. Inst. Henri Poincaré B (2015) in press.

[6] P. Billingsley, Convergence of Probability Measures, second ed., in: Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1999.
[7] T. Bodineau, C. Toninelli, Activity phase transition for constrained dynamics, Comm. Math. Phys. 311 (2012) 357–396.
[8] V.I. Bogachev, Measure Theory. Vol. II, Springer Verlag, Berlin, 2007.
[9] B. Bollobás, Modern Graph Theory, in: Graduate Texts in Mathematics, vol. 184, Springer Verlag, New York, 1998.
[10] A. de La Fortelle, The large-deviation principle for Markov chains with continuous time, Problemy Peredachi Informatsii 37 (2) (2001) 40–61. (in Russian); translation in Probl. Inf. Transm. 37 no. 2, 120–139 (2001).
[11] R. Diestel, private communication.
[12] R. Diestel, Graph Theory, Springer, Heidelberg, 2010.
[13] M.D. Donsker, S.R.S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time, Comm. Pure Appl. Math. (I) 28 (1975) 1–47. (II) 28, 279–301 (1975); (III) 29, 389–461 (1976); (IV) 36, 183–212 (1983).
[14] A. Faggionato, D. Di Pietro, Gallavotti–Cohen-type symmetry related to cycle decompositions for Markov chains and biochemical applications, J. Stat. Phys. 143 (2011) 11–32.
[15] M. Hamann, Accessibility in transitive graphs. Preprint 2014. Available online at: http://arxiv.org/abs/1404.7677.
[16] G. Kesidis, J. Walrand, Relative entropy between Markov transition rate matrices, IEEE Trans. Inform. Theory 39 (3) (1993) 1056–1057.
[17] M. Kotani, T. Sunada, Large deviation and the tangent cone at infinity of a crystal lattice, Math. Z. 254 (2006) 837–870.
[18] S. Kusuoka, K. Kuwada, Y. Tamura, Large deviation for stochastic line integrals as L^p-currents, Probab. Theory Related Fields 147 (2010) 649–674.
[19] K. Kuwada, On large deviations for random currents induced from stochastic line integrals, Forum Math. 18 (2006) 639–676.
[20] J.L. Lebowitz, H. Spohn, A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics, J. Stat. Phys. 95 (1999) 333–365.
[21] C. Maes, K. Netočný, Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states, Europhys. Lett. 82 (2008) 30003.
[22] C. Maes, K. Netočný, B. Wynants, Steady state statistics of driven diffusions, Physica A 387 (2008) 2675–2689.
[23] C. Maes, K. Netočný, B. Wynants, Dynamical fluctuations for semi-Markov processes, J. Phys. A: Math. Theor. 42 (2009) 365002.
[24] R.E. Megginson, An Introduction to Banach Space Theory, Springer-Verlag, New York, 1998.
[25] J.R. Norris, Markov Chains, in: Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 1999.
[26] J. Schnakenberg, Network theory of microscopic and macroscopic behavior of master equation systems, Rev. Modern Phys. 48 (1976) 571–585.
[27] U. Seifert, Stochastic thermodynamics, fluctuation theorems, and molecular machines, Rep. Progr. Phys. 75 (2012) 126001.
[28] S.R.S. Varadhan, Stochastic Processes, in: Courant Lecture Notes in Mathematics, vol. 16, American Mathematical Society, 2007.