



**HAL**  
open science

# Metastability from the large deviations point of view: A $\Gamma$ -expansion of the level two large deviations rate functional of non-reversible finite-state Markov chains

Claudio Landim

## ► To cite this version:

Claudio Landim. Metastability from the large deviations point of view: A  $\Gamma$ -expansion of the level two large deviations rate functional of non-reversible finite-state Markov chains. *Stochastic Processes and their Applications*, 2023, 165, pp.275-315. 10.1016/j.spa.2023.09.001 . hal-04284974

**HAL Id: hal-04284974**

**<https://hal.science/hal-04284974>**

Submitted on 14 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**METASTABILITY FROM THE LARGE DEVIATIONS POINT OF  
VIEW: A  $\Gamma$ -EXPANSION OF THE LEVEL TWO LARGE  
DEVIATIONS RATE FUNCTIONAL OF NON-REVERSIBLE  
FINITE-STATE MARKOV CHAINS**

C. LANDIM

ABSTRACT. Consider a sequence of continuous-time Markov chains  $(X_t^{(n)} : t \geq 0)$  evolving on a fixed finite state space  $V$ . Let  $\mathcal{J}_n$  be the level two large deviations rate functional for  $X_t^{(n)}$ , as  $t \rightarrow \infty$ . Under a hypothesis on the jump rates, we prove that  $\mathcal{J}_n$  can be written as  $\mathcal{J}_n = \mathcal{J}^{(0)} + \sum_{1 \leq p \leq q} (1/\theta_n^{(p)}) \mathcal{J}^{(p)}$  for some rate functionals  $\mathcal{J}^{(p)}$ . The weights  $\theta_n^{(p)}$  correspond to the time-scales at which the sequence of Markov chains  $X_t^{(n)}$  exhibit a metastable behavior, and the zero level sets of the rate functionals  $\mathcal{J}^{(p)}$  identify the metastable states.

1. INTRODUCTION

Fix a finite set  $V$  and consider a sequence  $(X_t^{(n)} : t \geq 0)$ ,  $n \geq 1$ , of  $V$ -valued, irreducible continuous-time Markov chains. Denote the jump rates by  $R_n : V \times V \rightarrow \mathbb{R}_+$ , and the generator by  $\mathcal{L}_n$ , so that

$$(\mathcal{L}_n f)(x) = \sum_{y \in V} R_n(x, y) \{ f(y) - f(x) \}, \quad f : V \rightarrow \mathbb{R}. \quad (1.1)$$

Let  $\pi_n$  be the unique stationary state.

Denote by  $\mathcal{P}(V)$  the space of probability measures on  $V$  endowed with the weak topology, and by  $L_t^{(n)}$  the empirical measure of the chain  $X_t^{(n)}$  defined as :

$$L_t^{(n)} := \frac{1}{t} \int_0^t \delta_{X_s^{(n)}} ds, \quad (1.2)$$

where  $\delta_x$ ,  $x \in V$ , represents the Dirac measure concentrated at  $x$ . Thus,  $L_t^{(n)}$  is a random element of  $\mathcal{P}(V)$  and  $L_t^{(n)}(V_0)$ ,  $V_0 \subset V$ , stands for the average amount of time the process  $X_t^{(n)}$  stays at  $V_0$  in the time interval  $[0, t]$ ,

As the Markov chain  $X_t^{(n)}$  is irreducible, by the ergodic theorem, for any starting point  $x \in V$ , as  $t \rightarrow \infty$ , the empirical measure  $L_t^{(n)}$  converges in probability to the stationary state  $\pi_n$ .

Donsker and Varadhan [10] proved the associated large deviations principle: for any  $x \in V$ ,  $\mu \in \mathcal{P}(V)$ ,

$$\mathbf{P}_x^n [L_t^{(n)} \sim \mu] \approx e^{-t \mathcal{J}_n(\mu)}, \quad \text{as } t \rightarrow \infty. \quad (1.3)$$

---

2010 *Mathematics Subject Classification.* Primary 60F10; 60J27; 60J45.

*Key words and phrases.* Metastability, Large deviations, Continuous-time Markov processes on discrete state spaces.

In this formula,  $\mathbf{P}_x = \mathbf{P}_x^n$ ,  $x \in V$ , represents the distribution of the process  $X_t^{(n)}$  starting from  $x$ , and  $\mathcal{J}_n: \mathcal{P}(V) \rightarrow [0, +\infty)$  be the level two large deviations rate functional given by

$$\mathcal{J}_n(\mu) := \sup_H J_H^{(n)}(\mu) := \sup_H - \int_V e^{-H} \mathcal{L}_n e^H d\mu, \quad (1.4)$$

where the supremum is carried over all functions  $H: V \rightarrow \mathbb{R}$ . A precise statement of (1.3) requires some notation and is postponed to the next section. The functional  $\mathcal{J}_n$  provides the cost for the empirical measure  $L_t^{(n)}$  to be close to  $\mu$  for a very large  $t$ . By Lemma A.8, as the process is irreducible,  $\mathcal{J}_n(\mu) = 0$  if, and only if,  $\mu = \pi_n$ ,

We examine in this article the behavior of the functionals  $\mathcal{J}_n$  as  $n \rightarrow \infty$  under some natural hypotheses on the jump rates  $R_n$ . Assume, initially, that the jump rates  $R_n(x, y)$  converge, as  $n \rightarrow \infty$ , to a limit represented by  $\mathbb{R}_0(x, y)$ :

$$\mathbb{R}_0(x, y) := \lim_n R_n(x, y) \in \mathbb{R}_+, \quad y \neq x \in V. \quad (1.5)$$

Denote by  $\mathbb{L}_0$  the generator associated to these rates and by  $\mathcal{J}^{(0)}$  the corresponding large deviations rate functional. By Proposition 2.1, as  $n \rightarrow \infty$ ,  $\mathcal{J}_n(\mu)$  converges to  $\mathcal{J}^{(0)}(\mu)$  for all  $\mu \in \mathcal{P}(V)$ .

If the Markov chain  $\mathbb{X}_t$  induced by the jump rates  $\mathbb{R}_0$  has only one closed irreducible class, the asymptotic analysis of the functionals  $\mathcal{J}_n$  ends with Proposition 2.1. In contrast, if  $\mathbb{X}_t$  has more than one closed irreducible class a finer description of  $\mathcal{J}_n$  is possible.

Denote by  $\mathcal{V}_1, \dots, \mathcal{V}_n$ ,  $n \geq 2$ , the closed irreducible classes of  $\mathbb{X}_t$ . Let  $\pi_j^\sharp$  be the stationary state supported in  $\mathcal{V}_j$ . By Lemma A.8,  $\mathcal{J}^{(0)}$  vanishes at any convex combination of the measures  $\pi_j^\sharp$ . Since, by Proposition 2.1,  $\mathcal{J}_n(\mu)$  converges to  $\mathcal{J}^{(0)}(\mu)$  it is natural to consider the sequence  $\theta_n \mathcal{J}_n(\mu)$ , for some  $\theta_n \rightarrow \infty$  and a convex combination  $\mu = \sum_j \omega_j \pi_j^\sharp$  of the measures  $\pi_j^\sharp$ , longing to obtain a non-trivial limit.

To find the correct sequence  $\theta_n$ , remark that, by (1.4),  $\theta_n \mathcal{J}_n$  represents the large deviations rate functional of the Markov chain induced by the generator  $\theta_n \mathcal{L}_n$ , that is, the rate functional of the Markov chain  $X_t^{(n)}$  observed at the time scale  $\theta_n$ :  $X_t^{1,n} := X_{\theta_n t}^{(n)}$ .

Denote by  $\beta_{n,j}$  the transition time from  $\mathcal{V}_j$  to  $\cup_{k \neq j} \mathcal{V}_k$ , this is the mean time for the process  $X_t^{(n)}$  to hit  $\cup_{k \neq j} \mathcal{V}_k$  when it starts from  $\mathcal{V}_j$ . For the sake of the argument, assume that  $\beta_n = \beta_{n,1}$ .

Fix a time-scale  $\theta_n$  such that  $\theta_n \rightarrow \infty$ ,  $\theta_n/\beta_n \rightarrow 0$ . Denote this last relation by  $\theta_n \prec \beta_n$  or  $\beta_n \succ \theta_n$ . As the transition time from  $\mathcal{V}_j$  to  $\cup_{k \neq j} \mathcal{V}_k$  is of order  $\beta_n$  and  $\beta_n \succ \theta_n$ , in the time-scale  $\theta_n$  starting from  $\mathcal{V}_j$  the chain  $X_t^{(n)}$  does not visit the set  $\cup_{k \neq j} \mathcal{V}_k$ . Therefore, the cost for keeping the process at  $\mathcal{V}_j$  should vanish, and one expects  $\theta_n \mathcal{J}_n(\pi_1^\sharp) \rightarrow 0$ . Actually, as  $\beta_{n,j} \succ \theta_n$  for all  $j$ , the same conclusion should hold for all measures  $\pi_j^\sharp$ , and to derive a non-trivial limit for  $\theta_n \mathcal{J}_n$  one has to observe the chain  $X_t^{(n)}$  in a time-scale at least of the order  $\beta_n$ .

In the time scale  $\beta_n$ , starting from  $\mathcal{V}_j$  the process visits  $\cup_{k \neq j} \mathcal{V}_k$ . There is, in consequence, a positive cost to maintain it at  $\mathcal{V}_j$  and  $\theta_n \mathcal{J}_n(\pi_1^\sharp)$  should converge to a positive limit. If all sequence  $\beta_{n,j}$  are of the same order, this completes the description of  $\mathcal{J}_n$ . Otherwise, one has to go to longer time-scales.

The main result of this article, Theorem 2.3, presents the time-scales  $1 \prec \theta_n^{(1)} \prec \dots \prec \theta_n^{(q)}$  and functionals  $\mathcal{J}^{(1)}, \dots, \mathcal{J}^{(q)}$  such that

$$\theta_n^{(p)} \mathcal{J}_n \longrightarrow \mathcal{J}^{(p)}, \quad 1 \leq p \leq q.$$

This result permits to write the functional  $\mathcal{J}_n$  as the expansion

$$\mathcal{J}_n = \mathcal{J}^{(0)} + \sum_{p=1}^q \frac{1}{\theta_n^{(p)}} \mathcal{J}^{(p)}. \quad (1.6)$$

The weights  $\theta_n^{(p)}$  correspond to the time-scales at which the sequence of Markov chains  $X_t^{(n)}$  exhibit a metastable behavior, and the zero level sets of the rate functionals  $\mathcal{J}^{(p)}$  identify the metastable states.

The proof of Theorem 2.3 relies on [4, 23] where the metastable behavior of the sequence  $X_t^{(n)}$  has been investigated. The expansion (1.6) has been derived for reversible diffusions in [12] and for reversible finite state Markov chains in [5]. It should be a universal property of Markov chains and should hold for dynamics whose state space depend on  $n$  and which exhibit a metastable behavior at different time-scales. This includes, among others models, random walks and diffusions in potential fields [7, 19–22, 24, 25, 28, 30], condensing zero-range processes [1, 3, 16, 27, 29], inclusion processes [6, 9, 13–15].

We believe that the argument proposed here to derive the expansion of the large deviations rate functional can be adapted to cover these dynamics.

## 2. NOTATION AND RESULTS

We present in this section the main result of the article. Consider a sequence  $(X_t^{(n)} : t \geq 0)$ ,  $n \geq 1$ , of  $V$ -valued, irreducible continuous-time Markov chains whose generator is given by (1.1).

Denote by  $D(\mathbb{R}_+, W)$ ,  $W$  a finite set, the space of right-continuous functions  $\mathfrak{x} : \mathbb{R}_+ \rightarrow W$  with left-limits endowed with the Skorohod topology and the associated Borel  $\sigma$ -algebra. Let  $\mathbf{P}_x = \mathbf{P}_x^n$ ,  $x \in V$ , be the distribution of the process  $X_t^{(n)}$  starting from  $x$ . This is the probability measure on the path space  $D(\mathbb{R}_+, V)$  induced by the Markov chain  $X_t^{(n)}$  starting from  $x$ . Expectation with respect to  $\mathbf{P}_x$  is represented by  $\mathbf{E}_x$ .

Recall the definition of the empirical measure  $L_t^{(n)}$  introduced in (1.2). Donsker and Varadhan [10] proved a large deviations principle for the empirical measure  $L_t^{(n)}$ . More precisely, they showed that for any subset  $\mathcal{A}$  of  $\mathcal{P}(V)$ ,

$$\begin{aligned} - \inf_{\mu \in \mathcal{A}^\circ} \mathcal{J}_n(\mu) &\leq \liminf_{t \rightarrow \infty} \inf_{x \in V} \frac{1}{t} \ln \mathbf{P}_x^{(n)} [L_t^{(n)} \in \mathcal{A}] \\ &\leq \limsup_{t \rightarrow \infty} \sup_{x \in V} \frac{1}{t} \ln \mathbf{P}_x^{(n)} [L_t^{(n)} \in \mathcal{A}] \leq - \inf_{\mu \in \bar{\mathcal{A}}} \mathcal{J}_n(\mu). \end{aligned} \quad (2.1)$$

In this formula,  $\mathcal{A}^\circ$ ,  $\bar{\mathcal{A}}$  represent the interior, closure of  $\mathcal{A}$ , respectively, and  $\mathcal{J}_n$  is the large deviations rate functional introduced in (1.4).

We examine in this article the asymptotic behavior of the rate functional  $\mathcal{J}_n$ . In the context of large deviations, the appropriate notion of convergence is the  $\Gamma$ -convergence defined as follows. We refer to [8] for an overview on this subject.

Fix a Polish space  $\mathcal{X}$  and a sequence  $(U_n : n \in \mathbb{N})$  of functionals on  $\mathcal{X}$ ,  $U_n : \mathcal{X} \rightarrow [0, +\infty]$ . The sequence  $U_n$   $\Gamma$ -converges to the functional  $U : \mathcal{X} \rightarrow [0, +\infty]$  if and only if the two following conditions are met:

- (i)  $\Gamma$ -*liminf*. The functional  $U$  is a  $\Gamma$ -liminf for the sequence  $U_n$ : For each  $x \in \mathcal{X}$  and each sequence  $x_n \rightarrow x$ , we have that  $\liminf_n U_n(x_n) \geq U(x)$ .
- (ii)  $\Gamma$ -*limsup*. The functional  $U$  is a  $\Gamma$ -limsup for the sequence  $U_n$ : For each  $x \in \mathcal{X}$  there exists a sequence  $x_n \rightarrow x$  such that

$$\limsup_{n \rightarrow \infty} U_n(x_n) \leq U(x). \quad (2.2)$$

Recall that we denote by  $R_n(x, y)$  the jump rates of the Markov chain  $X_t^{(n)}$ . Assume that the rates converge, as  $n \rightarrow \infty$ , to a finite limit denoted by  $\mathbb{R}_0(x, y)$ , see (1.5), and that  $\mathbb{R}_0(x', y') > 0$  for some  $y' \neq x' \in V$ . The jump rates  $\mathbb{R}_0(x, y)$  induce a continuous-time Markov chain on  $V$ , denoted by  $(\mathbb{X}_t : t \geq 0)$ , which, of course, may be reducible. Denote by  $\mathbb{L}^{(0)}$  its generator and by  $\mathcal{J}^{(0)} : \mathcal{P}(V) \rightarrow \mathbb{R}_+$  the associated occupation-time large deviations rate functional, given by

$$\mathcal{J}^{(0)}(\mu) = \sup_H - \sum_{x \in V} e^{-H(x)} [(\mathbb{L}^{(0)} e^H)(x)] \mu(x), \quad (2.3)$$

where the supremum is carried over all functions  $H : V \rightarrow \mathbb{R}$ . Next result is proved in Section 5.

**Proposition 2.1.** *The sequence of functionals  $\mathcal{J}_n$   $\Gamma$ -converges to  $\mathcal{J}^{(0)}$ .*

Assume from now on that

$$\text{the Markov chain } \mathbb{X}_t \text{ has more than one closed irreducible class.} \quad (2.4)$$

Under this hypothesis we may investigate further the asymptotic behavior of the rate functional  $\mathcal{J}_n$ .

**Main assumption.** To examine the convergence of  $\theta_n \mathcal{J}_n$  for some sequence  $\theta_n \rightarrow \infty$ , we introduce a natural hypothesis on the jump rates proposed in [4] and adopted in [5, 11, 23].

For two sequences of positive real numbers  $(\alpha_n : n \geq 1)$ ,  $(\beta_n : n \geq 1)$ ,  $\alpha_n \prec \beta_n$  or  $\beta_n \succ \alpha_n$  means that  $\lim_{n \rightarrow \infty} \alpha_n / \beta_n = 0$ . Similarly,  $\alpha_n \preceq \beta_n$  or  $\beta_n \succeq \alpha_n$  indicates that either  $\alpha_n \prec \beta_n$  or  $\alpha_n / \beta_n$  converges to a positive real number  $a \in (0, \infty)$ .

Two sequences of positive real numbers  $(\alpha_n : n \geq 1)$ ,  $(\beta_n : n \geq 1)$  are said to be *comparable* if  $\alpha_n \prec \beta_n$ ,  $\beta_n \prec \alpha_n$  or  $\alpha_n / \beta_n \rightarrow a \in (0, \infty)$ . This condition excludes the possibility that  $\liminf_n \alpha_n / \beta_n \neq \limsup_n \alpha_n / \beta_n$ .

A set of sequences  $(\alpha_n^{\mathbf{u}} : n \geq 1)$ ,  $\mathbf{u} \in \mathfrak{A}$ , of positive real numbers, indexed by some finite set  $\mathfrak{A}$ , is said to be comparable if for all  $\mathbf{u}, \mathbf{v} \in \mathfrak{A}$  the sequence  $(\alpha_n^{\mathbf{u}} : n \geq 1)$ ,  $(\alpha_n^{\mathbf{v}} : n \geq 1)$  are comparable.

Denote by  $E \subset \{(x, y) \in V \times V : y \neq x\}$  a set of directed edges, and assume that for all  $n \geq 1$ ,

$$R_n(x, y) > 0 \text{ if, and only if, } (x, y) \in E. \quad (2.5)$$

Let  $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ , and  $\Sigma_m$ ,  $m \geq 1$ , be the set of functions  $k : E \rightarrow \mathbb{Z}_+$  such that  $\sum_{(x, y) \in E} k(x, y) = m$ . We assume, hereafter, that for every  $m \geq 1$  the set of sequences

$$\left( \prod_{(x, y) \in E} R_n(x, y)^{k(x, y)} : n \geq 1 \right), \quad k \in \Sigma_m, \quad (2.6)$$

is comparable.

As observed in [4] (see Remark 2.2 in [5]), assumption (2.6) is fulfilled by all statistical mechanics models which evolve on a fixed finite state space and whose metastable behaviour has been derived.

**Tree decomposition.** Under the assumptions (2.4), (2.5) and (2.6), [4, 23] constructed a rooted tree which describes the behaviour of the Markov chain  $X_t^{(n)}$  at all different time-scales. We refer to [5] for a clear presentation of the construction as well as a simple example, and recall here the main ideas.

Denote by  $q + 1 \geq 2$  the number of generations of the tree. The elements of the  $p$ -th generation form a partition of  $V$ , and are represented by  $\mathcal{W}_1^{(p)}, \dots, \mathcal{W}_{\mathbf{m}_p}^{(p)}, \Omega_p$  for some finite increasing sequence  $1 = \mathbf{m}_1 \leq \dots \leq \mathbf{m}_{q+1}$ . The set  $\Omega_p$  may be empty while the sets  $\mathcal{W}_j^{(p)}$  are all non-empty. As  $\mathbf{m}_p \geq 1$ , each generation has at least one element. Here is a list of the main properties of the tree:

- (1.a) Each generation of the tree forms a partition of  $V$ ;
- (1.b) The root, or 0-th generation, is the set  $V$ . The first generation has one or two elements depending on whether  $\Omega_1$  is empty or not. If  $\Omega_1 = \emptyset$ , it has one element, the set  $\mathcal{W}_1^{(1)} = V$ . If  $\Omega_1 \neq \emptyset$ , it has two elements, the sets  $\mathcal{W}_1^{(1)}$  and  $\Omega_1 = (\mathcal{W}_1^{(1)})^c$ .
- (1.c) Each child of a vertex is a subset of its parent: For each  $0 \leq p \leq q$ ,  $1 \leq j \leq \mathbf{m}_{p+1}$ , either  $\mathcal{W}_j^{(p+1)} \subset \mathcal{W}_k^{(p)}$  for some  $1 \leq k \leq \mathbf{m}_p$  or  $\mathcal{W}_j^{(p+1)} \subset \Omega_p$ . Moreover,  $\Omega_{p+1} \subset \Omega_p$ ;
- (1.d) According to the notation, the number of elements of generation  $p$  is equal to  $\mathbf{m}_p + \mathbf{1}_{\Omega_p \neq \emptyset}$ , where  $\mathbf{1}_{\mathcal{A} \neq \emptyset}$  is equal to 1 if  $\mathcal{A}$  is not empty and 0 otherwise. Starting from the first generation, the number of descendants of a generation strictly increases: for  $1 \leq p \leq q$ ,  $\mathbf{m}_p + \mathbf{1}_{\Omega_p \neq \emptyset} < \mathbf{m}_{p+1} + \mathbf{1}_{\Omega_{p+1} \neq \emptyset}$ .

**Construction of the tree.** We describe in this subsection the details of the construction of the tree. It is formed from the leaves to the root. The leaves  $\mathcal{V}_1^{(1)}, \dots, \mathcal{V}_{\mathbf{n}_1}^{(1)}$  are the closed irreducible classes of the Markov chain  $\mathbb{X}_t$  introduced in the previous section, and  $\Delta_1$  the set of transient states. The sets  $\mathcal{V}_j^{(1)}$  were represented there by  $\mathcal{V}_j$ , a notation frequently adopted below. In view of the definition of the sets  $\mathcal{V}_j^{(1)}$ ,  $\mathbf{n}_1$  corresponds to the number of closed irreducible classes of the process  $\mathbb{X}_t$ , which we assumed in (2.4) to be larger than or equal to 2.

We turn to the construction of the parents of the leaves. This procedure will be repeated recursively to define all generations from the leaves to the root. Denote by  $\Phi_1 : V \rightarrow S_1$  the projection defined by

$$\Phi_1(\cdot) = \sum_{j \in S_1} j \chi_{\mathcal{V}_j^{(1)}}(\cdot),$$

where  $\chi_{\mathcal{A}}$  stands for the indicator function of the set  $\mathcal{A}$ . Hence,  $\Phi_1$  projects to 0 all elements of  $\Delta_1$  and to  $j$  the ones of  $\mathcal{V}_j^{(1)}$ .

It follows from the main results of [18, 23] and [5, Lemmata 4.7] that there exist a time-scale  $\theta_n^{(1)} \succ 1$  and a  $S_1$ -valued Markov chain  $\mathbb{X}_t^{(1)}$  (note that this process does not take the value 0), such that the finite-dimensional distributions of  $\Phi_1(X_{t\theta_n^{(1)}}^{(n)})$

converge to those of  $\mathbb{X}_1^{(1)}$ :

$$\Phi_1(X_{t\theta_n^{(1)}}^{(n)}) \xrightarrow{\text{f.d.d.}} \mathbb{X}_t^{(1)}. \quad (2.7)$$

Denote by  $\mathfrak{R}_1^{(1)}, \dots, \mathfrak{R}_{n_2}^{(1)}$  the recurrent classes of the  $S_1$ -valued Markov chain  $\mathbb{X}_t^{(1)}$ , and by  $\mathfrak{T}_1$  the transient states. Let  $\mathfrak{R}^{(1)} = \cup_j \mathfrak{R}_j^{(1)}$ , and observe that  $\{\mathfrak{R}_1^{(1)}, \dots, \mathfrak{R}_{n_2}^{(1)}, \mathfrak{T}_1\}$  forms a partition of the set  $S_1$ . This partition of  $S_1$  induces a new partition of  $V$ . Let

$$\mathcal{V}_m^{(2)} := \bigcup_{j \in \mathfrak{R}_m^{(1)}} \mathcal{V}_j^{(1)}, \quad \mathcal{T}^{(1)} := \bigcup_{j \in \mathfrak{T}_1} \mathcal{V}_j^{(1)}, \quad m \in S_2 := \{1, \dots, n_2\}, \quad (2.8)$$

so that  $V = \Delta_2 \cup \mathcal{V}^{(2)}$ , where

$$\mathcal{V}^{(2)} = \bigcup_{m \in S_2} \mathcal{V}_m^{(2)}, \quad \Delta_2 := \Delta_1 \cup \mathcal{T}^{(2)}.$$

It is shown in [23] that the Markov chain  $\mathbb{X}_t^{(1)}$  is non-degenerate in the sense that there exists at least one edge  $(j, k)$ ,  $k \neq j \in S_1$ , such that  $r^{(1)}(j, k) > 0$ , where  $r^{(1)}(\cdot, \cdot)$  represents the jump rates of the Markov chain  $\mathbb{X}_t^{(1)}$ . In particular, either  $j$  is a transient state or  $j$  and  $k$  belong to the same closed irreducible class. Therefore, the number of recurrent classes ( $n_2$ ) is strictly smaller than the number of  $S_1$  elements ( $n_1$ ):  $n_2 < n_1$ . Since, on the other hand,  $\Delta_2 \supset \Delta_1$ , the number of leaves' parents (the generation  $q-1$  in the previous subsection) is strictly smaller than the one of leaves (the generation  $q$ ).

In conclusion, from the partition  $\mathcal{V}_1^{(1)}, \dots, \mathcal{V}_{n_1}^{(1)}, \Delta_1$ , the theory presented in [23] produced a time-scale  $\theta_n^{(1)} \succ 1$ , a  $S_1$ -valued Markov chain  $\mathbb{X}_t^{(1)}$ , and a coarser partition  $\mathcal{V}_1^{(2)}, \dots, \mathcal{V}_{n_2}^{(2)}, \Delta_2$ . The construction of the tree proceeds by recurrence.

Assume that, for some  $p > 1$ , the recursion has produced

- (a) Time scales  $1 < \theta_n^{(1)} < \dots < \theta_n^{(p-1)}$ ;
- (b)  $S_q$ -valued Markov chains  $\mathbb{X}_t^{(q)}$ ,  $1 \leq q < p$ , where  $S_q = \{1, \dots, n_q\}$ ;
- (c) Partitions  $\mathcal{V}_1^{(r)}, \dots, \mathcal{V}_{n_r}^{(r)}, \Delta_r$ ,  $1 \leq r \leq p$

satisfying (2.7), (2.8) (with the obvious modifications which appear in (2.9), (2.10)). Assume, furthermore, that  $n_p > 1$ . Then, by [18, 23] and [5, Lemmata 5.6], there exist a time-scale  $\theta_n^{(p)} \succ \theta_n^{(p-1)}$  and a  $S_p$ -valued Markov chain  $\mathbb{X}_t^{(p)}$  such that the finite-dimensional distributions of  $\Phi_p(X_{t\theta_n^{(p)}}^{(n)})$  converge to those of  $\mathbb{X}_1^{(p)}$ :

$$\Phi_p(X_{t\theta_n^{(p)}}^{(n)}) \xrightarrow{\text{f.d.d.}} \mathbb{X}_t^{(p)}. \quad (2.9)$$

In this formula,  $\Phi_p : V \rightarrow S_p$  represents the projection defined by

$$\Phi_p(\cdot) = \sum_{j \in S_p} j \chi_{\mathcal{V}_j^{(p)}}(\cdot).$$

Denote by  $\mathfrak{R}_1^{(p)}, \dots, \mathfrak{R}_{n_{p+1}}^{(p)}$  the recurrent classes of the  $S_p$ -valued Markov chain  $\mathbb{X}_t^{(p)}$ , and by  $\mathfrak{T}_p$  the transient states. Let  $\mathfrak{R}^{(p)} = \cup_j \mathfrak{R}_j^{(p)}$ , and observe that  $\{\mathfrak{R}_1^{(p)}, \dots, \mathfrak{R}_{n_{p+1}}^{(p)}, \mathfrak{T}_p\}$  forms a partition of the set  $S_p$ . This partition of  $S_p$  induces a new

partition of  $V$ . Let

$$\mathcal{V}_m^{(p+1)} := \bigcup_{j \in \mathfrak{R}_m^{(p)}} \mathcal{V}_j^{(p)}, \quad \mathcal{J}^{(p)} := \bigcup_{j \in \mathfrak{I}_p} \mathcal{V}_j^{(p)}, \quad m \in S_{p+1} := \{1, \dots, \mathfrak{n}_{p+1}\}, \quad (2.10)$$

so that  $V = \Delta_{p+1} \cup \mathcal{V}^{(p+1)}$ , where

$$\mathcal{V}^{(p+1)} = \bigcup_{m \in S_{p+1}} \mathcal{V}_m^{(p+1)}, \quad \Delta_{p+1} := \Delta_p \cup \mathcal{J}^{(p+1)}.$$

As above, it is shown in [23] that the Markov chain  $\mathbb{X}_t^{(p)}$  is non-degenerate so that  $\mathfrak{n}_{p+1} < \mathfrak{n}_p$ . The induction can proceed if  $\mathfrak{n}_{p+1} > 1$ , otherwise it ends. Denote by  $\mathfrak{q}$  the first integer  $r$  such that  $\mathfrak{n}_{r+1} = 1$ , (equivalently, the first  $r$  such that the Markov chain  $\mathbb{X}_t^{(r)}$  has only one recurrent class). At this point the iteration stops and the partition of  $V$  produced is  $\{\mathcal{V}_1^{(\mathfrak{q}+1)}, \Delta_{\mathfrak{q}+1}\}$  which may have one or two elements, depending on whether  $\Delta_{\mathfrak{q}+1}$  is empty or not.

To recover the tree presented in the previous subsection, add a final partition equal to  $V$  which will identified to the root of the tree, and for  $1 \leq p \leq \mathfrak{q} + 1$ ,  $k \in S_{\mathfrak{q}+2-p} = \{1, \dots, \mathfrak{n}_{\mathfrak{q}+2-p}\}$ , set

$$\mathfrak{m}_p := \mathfrak{n}_{\mathfrak{q}+2-p}, \quad \mathcal{W}_k^{(p)} := \mathcal{V}_k^{(\mathfrak{q}+2-p)}, \quad \Omega_p = \Delta_{\mathfrak{q}+2-p}.$$

It is easy to check that conditions (1.a)–(1.d) are fulfilled.

**A set of measures.** We construct in this subsection a set of probability measures  $\pi_j^{(p)}$ ,  $1 \leq p \leq \mathfrak{q} + 1$ ,  $j \in S_p$ , on  $V$  which describe the evolution of the chain  $X_t^{(n)}$  and such that

$$\text{the support of } \pi_j^{(p)} \text{ is the set } \mathcal{V}_j^{(p)}. \quad (2.11)$$

We proceed by induction. Let  $\pi_j^{(1)}$ ,  $j \in S_1$ , be the probability measure on  $\mathcal{V}_j^{(1)}$  given by  $\pi_j^{(1)} = \pi_j^\sharp$ , where, recall,  $\pi_j^\sharp$  represents the stationary states of the Markov chain  $\mathbb{X}_t$  restricted to the closed irreducible class  $\mathcal{V}_j^{(1)} = \mathcal{V}_j$ . Clearly, condition (2.11) is fulfilled.

Fix  $1 \leq p \leq \mathfrak{q}$ , and assume that the probability measures  $\pi_j^{(p)}$ ,  $j \in S_p$ , has been defined and satisfy condition (2.11). Denote by  $M_m^{(p)}(\cdot)$ ,  $m \in S_{p+1}$ , the stationary state of the Markov chain  $\mathbb{X}_t^{(p)}$  restricted to the closed irreducible class  $\mathfrak{R}_m^{(p)}$ . The measure  $M_m^{(p)}$  is understood as a measure on  $S_p = \{1, \dots, \mathfrak{n}_p\}$  which vanishes on the complement of  $\mathfrak{R}_m^{(p)}$ . Let  $\pi_m^{(p+1)}$  be the probability measure on  $\mathcal{V}_m^{(p+1)}$  given by

$$\pi_m^{(p+1)}(x) := \sum_{j \in \mathfrak{R}_m^{(p)}} M_m^{(p)}(j) \pi_j^{(p)}(x), \quad x \in V. \quad (2.12)$$

Clearly, condition (2.11) holds, and the measure  $\pi_m^{(p+1)}$ ,  $1 \leq p \leq \mathfrak{q}$ ,  $m \in S_{p+1}$ , is a convex combination of the measures  $\pi_j^{(p)}$ ,  $j \in \mathfrak{R}_m^{(p)}$ . Moreover, by [5, Theorem 3.1 and Proposition 3.2], for all  $z \in \mathcal{V}_j^{(p)}$ ,

$$\lim_{n \rightarrow \infty} \frac{\pi_n(z)}{\pi_n(\mathcal{V}_j^{(p)})} = \pi_j^{(p)}(z) \in (0, 1], \quad \lim_{n \rightarrow \infty} \pi_n(\Delta_{\mathfrak{q}+1}) = 0. \quad (2.13)$$



By (2.12), the measures  $\pi_j^{(p)}$ ,  $2 \leq p \leq \mathfrak{q} + 1$ ,  $j \in S_p$ , are convex combinations of the measures  $\pi_k^{(1)}$ ,  $k \in S_1$ . By (2.13), for all  $x \in \mathcal{V}^{(\mathfrak{q}+1)}$ ,  $\lim_{n \rightarrow \infty} \pi_n(x)$  exists and belongs to  $(0, 1]$ . By (2.13), and since by (1.c)  $\Delta_p \subset \Delta_{p+1}$  for  $1 \leq p \leq \mathfrak{q}$ ,  $\lim_{n \rightarrow \infty} \pi_n(\Delta_p) = 0$  for all  $p$ .

**A complete description of the chain  $X_t^{(n)}$ .** The statement of next result requires some notation. Denote by  $H_{\mathcal{A}}$ ,  $H_{\mathcal{A}}^+$ ,  $\mathcal{A} \subset V$ , the hitting and return time of  $\mathcal{A}$ :

$$H_{\mathcal{A}} := \inf \{t > 0 : X_t^{(n)} \in \mathcal{A}\}, \quad H_{\mathcal{A}}^+ := \inf \{t > \tau_1 : X_t^{(n)} \in \mathcal{A}\}, \quad (2.14)$$

where  $\tau_1$  represents the time of the first jump of the chain  $X_t^{(n)}$ :  $\tau_1 = \inf \{t > 0 : X_t^{(n)} \neq X_0^{(n)}\}$ .

For  $1 \leq p \leq \mathfrak{q} + 1$ ,  $k \in S_p$ , let

$$\check{\mathcal{V}}_k^{(p)} := \bigcup_{j \in S_p \setminus \{k\}} \mathcal{V}_j^{(p)}.$$

Define  $\mathbf{a}^{(p-1)} : V \times S_p \rightarrow [0, 1]$  as follows. Fix  $j \in S_p$ . If  $x \notin \mathcal{V}^{(p)}$ , set

$$\mathbf{a}^{(p-1)}(x, j) := \lim_{n \rightarrow \infty} \mathbf{P}_x^n [H_{\check{\mathcal{V}}_j^{(p)}} < H_{\check{\mathcal{V}}_j^{(p)}}],$$

while, if  $x \in \mathcal{V}_k^{(p)}$  for  $k \in S_p$ , set  $\mathbf{a}^{(p-1)}(x, j) = \delta_{j,k}$ . For  $p = \mathfrak{q} + 1$ , as  $S_{\mathfrak{q}+1}$  is a singleton,  $\mathbf{a}^{(\mathfrak{q})}(x, 1) = 1$  for all  $x \in V$ .

Denote by  $p_t^{(n)}(x, y)$  the transition probability of the Markov chain  $X_t^{(n)}$ :

$$p_t^{(n)}(x, y) := \mathbf{P}_x^n [X_t = y], \quad x, y \in V, \quad t > 0.$$

Next result is [5, Theorem 3.1 and Proposition 3.2].

**Theorem 2.2.** *Under the hypotheses (2.5) and (2.6), for each  $1 \leq p \leq \mathfrak{q}$ ,  $t > 0$ ,  $x \in V$ ,*

$$\lim_{n \rightarrow \infty} p_{t\theta_n^{(p)}}^{(n)}(x, \cdot) = \sum_{j \in S_p} \omega_t^{(p)}(x, j) \pi_j^{(p)}(\cdot), \quad (2.15)$$

where

$$\omega_t^{(p)}(x, j) = \sum_{k \in S_p} \mathbf{a}^{(p-1)}(x, k) p_t^{(p)}(k, j),$$

and  $p_t^{(p)}(k, j)$  is the transition matrix of the Markov chain  $\mathbb{X}_t^{(p)}$ . Moreover,

(3.a) *Let  $\theta_n^{(0)} = 1$ ,  $\theta_n^{(\mathfrak{q}+1)} = +\infty$  for all  $n \geq 1$ . For each  $1 \leq p \leq \mathfrak{q} + 1$ , sequence  $(\beta_n : n \geq 1)$  such that  $\theta_n^{(p-1)} \prec \beta_n \prec \theta_n^{(p)}$ , and  $x \in V$ ,*

$$\lim_{n \rightarrow \infty} p_{\beta_n}^{(n)}(x, \cdot) = \sum_{j \in S_p} \mathbf{a}^{(p-1)}(x, j) \pi_j^{(p)}(\cdot).$$

(3.b) *For all  $1 \leq p \leq \mathfrak{q}$ ,  $1 \leq j \leq \mathfrak{n}_p$ ,  $x \in V$ ,*

$$\lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} p_{t\theta_n^{(p)}}^{(n)}(x, \cdot) = \sum_{m \in S_{p+1}} \mathbf{a}^{(p)}(x, m) \pi_m^{(p+1)}(\cdot).$$

Equation (2.15) and properties (3.a), (3.b) describe the behavior of the Markov chain  $X_t^{(n)}$  in all time-scales. By (2.15), for instance, starting from  $x$ , as  $n \rightarrow \infty$ , the distribution of  $X_{t\theta_n^{(p)}}^{(n)}$  is a convex combination of the measures  $\pi_j^{(p)}$ . The weights  $\omega_t^{(p)}(x, j)$  have a simple interpretation:  $\omega_t^{(p)}(x, j)$  is equal to the probability that

starting from  $x$  the chain reaches the set  $\mathcal{V}^{(p)}$  at  $\mathcal{V}_k^{(p)}$  times the probability that the Markov chain  $\mathbb{X}_t^{(p)}$  starting from  $k$  is at  $j$  at time  $t$ .

Clearly, by (2.15), for all  $1 \leq p \leq \mathfrak{q}$ ,  $j \in S_p$ ,  $x \in V$ ,

$$\lim_{t \rightarrow 0} \lim_{n \rightarrow \infty} p_{t\theta_n^{(p)}}^{(n)}(x, \cdot) = \sum_{j \in S_p} \mathbf{a}^{(p-1)}(x, j) \pi_j^{(p)}(\cdot).$$

**The  $\Gamma$ -expansion of the rate functional  $\mathcal{J}_n$ .** We are now in a position to state the main result of this article. Denote by  $\mathcal{P}(S_p)$ ,  $1 \leq p \leq \mathfrak{q}$ , the set of probability measures on  $S_p$  and by  $\mathbb{L}^{(p)}$  the generator of the  $S_p$ -valued Markov chain  $\mathbb{X}_t^{(p)}$ . Let  $\mathbb{I}^{(p)}: \mathcal{P}(S_p) \rightarrow [0, +\infty)$  be the level two large deviations rate functional of  $\mathbb{X}_t^{(p)}$  given by

$$\mathbb{I}^{(p)}(\omega) := \sup_{\mathbf{h}} - \sum_{j \in S_p} \omega_j e^{-\mathbf{h}(j)} (\mathbb{L}^{(p)} e^{\mathbf{h}})(j), \quad (2.16)$$

where the supremum is carried over all functions  $\mathbf{h}: S_p \rightarrow \mathbb{R}$ . Denote by  $\mathcal{J}^{(p)}: \mathcal{P}(V) \rightarrow [0, +\infty]$  the functional given by

$$\mathcal{J}^{(p)}(\mu) := \begin{cases} \mathbb{I}^{(p)}(\omega) & \text{if } \mu = \sum_{j \in S_p} \omega_j \pi_j^{(p)} \text{ for } \omega \in \mathcal{P}(S_p), \\ +\infty & \text{otherwise.} \end{cases} \quad (2.17)$$

The main result of the article reads as follows.

**Theorem 2.3.** *For each  $1 \leq p \leq \mathfrak{q}$ , the functional  $\theta_n^{(p)} \mathcal{J}_n$   $\Gamma$ -converges to  $\mathcal{J}^{(p)}$ .*

This theorem provides an expansion of the large deviations rate function  $\mathcal{J}_n$  which can be written as

$$\mathcal{J}_n = \mathcal{J}^{(0)} + \sum_{p=1}^{\mathfrak{q}} \frac{1}{\theta_n^{(p)}} \mathcal{J}^{(p)}. \quad (2.18)$$

Therefore, the rate function  $\mathcal{J}_n$  encodes all the characteristics of the metastable behavior of the chain  $X_t^{(n)}$ . The time-scales  $\theta_n^{(p)}$  appear as the weights of the expansion, and, by (2.17), the meta-stable states  $\pi_j^{(p)}$ ,  $j \in S_p$ , generate the space where the rate functional  $\mathcal{J}^{(p)}(\mu)$  is finite.

Next result is a simple consequence of the level two large deviations principle (2.1) and the  $\Gamma$ -convergence stated in the previous theorem and in Proposition 2.1. (cf. Corollary 4.3 in [26]).

**Corollary 2.4.** *Fix  $0 \leq p \leq \mathfrak{q}$  and recall that  $\theta_n^{(0)} = 1$ . For every closed subset  $F$  and open subset  $G$  of  $\mathcal{P}(V)$ ,*

$$\begin{aligned} \limsup_{n \rightarrow \infty} \limsup_{t \rightarrow \infty} \frac{\theta_n^{(p)}}{t} \sup_{x \in V} \log \mathbf{P}_x^n \left[ \frac{1}{t} \int_0^t \delta_{X_s^n} ds \in F \right] &\leq - \inf_{\mu \in F} \mathcal{J}^{(p)}(\mu), \\ \liminf_{n \rightarrow \infty} \liminf_{t \rightarrow \infty} \frac{\theta_n^{(p)}}{t} \inf_{x \in V} \log \mathbf{P}_x^n \left[ \frac{1}{t} \int_0^t \delta_{X_s^n} ds \in G \right] &\geq - \inf_{\mu \in G} \mathcal{J}^{(p)}(\mu). \end{aligned}$$

**Organisation of the paper.** The article is organised as follows. In Section 3, we obtain some estimates on the jump rates. This is a technical section which can be skipped in a first reading. In Section 4, we prove the  $\Gamma$  – lim sup for the sequence of large deviations rate functionals associated to the trace process. Proposition 2.1 and Theorem 2.3 are proved in Section 5.

In the appendices we present general results on finite state Markov chains needed in the proof of Theorem 2.3 and which do not require assumption (2.6). In Appendix A we derive some properties of level two large deviations rate functionals. This leads us to introduce reflected and tilted dynamics. In Appendix B we investigate the convergence of these functionals, and in Appendix C the relation between the trace process and the rate functionals. Throughout the article we assume the reader to be familiar with the results presented in the appendices.

### 3. THE JUMP RATES

In this section, we state some estimates of the jump rates of the trace process on the sets  $\mathcal{V}^{(p)}$  needed in the next sections. We assume that the reader is familiar with the notation and results presented in the appendix.

Fix  $1 \leq p \leq \mathfrak{q}$ , and denote by  $\{Y_t^{n,p} : t \geq 0\}$  the trace of  $\{X_t^{(n)} : t \geq 0\}$  on  $\mathcal{V}^{(p)}$ , and by  $R_n^{(p)} : \mathcal{V}^{(p)} \times \mathcal{V}^{(p)} \rightarrow \mathbb{R}_+$  its jump rates. By equation (2.5) in [17],

$$R_n^{(p)}(x, y) = \lambda_n(x) \mathbf{P}_x^n [H_y = H_{\mathcal{V}^{(p)}}^+] , \quad x, y \in \mathcal{V}^{(p)}, x \neq y . \quad (3.1)$$

Let  $r_n^{(p)}(i, j)$ ,  $j \neq i \in S_p$ , be the mean rate at which the trace process  $Y_t^{n,p}$  jumps from  $\mathcal{V}_i^{(p)}$  to  $\mathcal{V}_j^{(p)}$ :

$$r_n^{(p)}(i, j) := \frac{1}{\pi_n(\mathcal{V}_i^{(p)})} \sum_{x \in \mathcal{V}_i^{(p)}} \pi_n(x) \sum_{y \in \mathcal{V}_j^{(p)}} R_n^{(p)}(x, y) . \quad (3.2)$$

By [23, Theorem 2.7 and 2.12], the sequences  $\theta_n^{(p)} r_n^{(p)}(i, j)$  converge for all  $i \neq j \in S_p$ . Denote the limits by  $r^{(p)}(i, j)$ :

$$r^{(p)}(i, j) := \lim_{n \rightarrow \infty} \theta_n^{(p)} r_n^{(p)}(i, j) \in \mathbb{R}_+ . \quad (3.3)$$

Recall from (A.10) the definition of the reflection of a Markov process on a subset of its state space.

**Lemma 3.1.** *For all  $n \geq 1$ ,  $1 \leq p \leq \mathfrak{q}$ ,  $j \in S_p$ , the trace process  $Y_t^{n,p}$  reflected at  $\mathcal{V}_j^{(p)}$  is irreducible,*

*Proof.* Fix  $1 \leq p \leq \mathfrak{q}$ ,  $j \in S_p$ ,  $x, y \in \mathcal{V}_j^{(p)}$ . We have to prove that there exists a path  $(x = x_0, x_1, \dots, x_\ell = y)$  such that  $x_i \in \mathcal{V}_j^{(p)}$ ,  $R_n^{(p)}(x_i, x_{i+1}) > 0$  for all  $0 \leq i < \ell$ ,  $n \geq 1$ .

By Propositions 6.1 and 6.3 in [2], the trace process  $Y_t^{n,p}$  is an irreducible,  $\mathcal{V}^{(p)}$ -valued continuous-time Markov chain. Fix  $j \in S_p$  and denote by  $Y_t^{n,p,j}$  the process  $Y_t^{n,p}$  reflected at  $\mathcal{V}_j^{(p)}$ .

The proof is by induction on  $p$ . Fix  $p = 1$  and consider the reflected process  $Y_t^{n,1,j}$  for  $j \in S_1$ . By definition of  $\mathcal{V}^{(1)}$ , the set  $\mathcal{V}_j^{(1)}$  is a closed irreducible class for the chain  $\mathbb{X}_t$ . Therefore, for all  $x \neq y \in \mathcal{V}_j^{(1)}$ , there exists a path  $(x = x_0, x_1, \dots, x_\ell = y)$  such that  $x_i \in \mathcal{V}_j^{(1)}$ ,  $\mathbb{R}_0(x_i, x_{i+1}) > 0$ ,  $0 \leq i < \ell$ . By assumptions (1.5), (2.5), for

all  $n \geq 1$ ,  $R_n(x_i, x_{i+1}) > 0$  as well, and by (C.3),  $R_n^{(1)}(x_i, x_{i+1}) > 0$ , completing the proof for  $p = 1$ .

Fix  $p > 1$ , and assume that the assertion of the lemma holds for  $1 \leq q < p$ . Consider the reflected process  $Y_t^{n,p,m}$  for  $m \in S_p$ , and fix  $y \neq x \in \mathcal{V}_m^{(p)}$ . By definition of  $\mathcal{V}_m^{(p)}$ , there exists a subset  $S_{p,m} \subset S_{p-1}$  such that  $\mathcal{V}_m^{(p)} = \cup_{j \in S_{p,m}} \mathcal{V}_j^{(p-1)}$ .

There are two cases. Assume first that  $x$  and  $y$  belong to the same set  $\mathcal{V}_j^{(p-1)}$ . By the induction assumption, there exists a path  $(x = x_0, x_1, \dots, x_\ell = y)$  such that  $x_i \in \mathcal{V}_j^{(p-1)}$ ,  $R_n^{(p-1)}(x_i, x_{i+1}) > 0$  for all  $0 \leq i < \ell$  and  $n \geq 1$ . By (C.3),  $R_n^{(p)}(x_i, x_{i+1}) \geq R_n^{(p-1)}(x_i, x_{i+1})$ , so that  $R_n^{(p)}(x_i, x_{i+1}) > 0$  for all  $0 \leq i < \ell$  and  $n \geq 1$ .

Assume now that  $x \in \mathcal{V}_j^{(p-1)}$  and  $y \in \mathcal{V}_k^{(p-1)}$  for  $k \neq j \in S_{p,m}$ . By construction of  $\mathcal{V}_m^{(p)}$ , there exists a sequence  $(j = j_0, j_1, \dots, j_r = k)$  such that  $j_a \in S_{p,m}$ ,  $r^{(p-1)}(j_a, j_{a+1}) > 0$ ,  $0 \leq a < r$ . To keep the proof simple assume that  $r^{(p-1)}(j, k) > 0$ . The reader will see that the proof in the general case is similar.

Since  $r^{(p-1)}(j, k) > 0$ , by (3.3), (3.2) and (2.13), there exists  $x' \in \mathcal{V}_j^{(p-1)}$ ,  $y' \in \mathcal{V}_k^{(p-1)}$  such that  $R_n^{(p-1)}(x', y') > 0$  for  $n$  sufficiently large. By (2.5) and (C.2),  $R_n^{(p-1)}(x', y') > 0$  for all  $n \geq 1$ . Hence, by (C.3),  $R_n^{(p)}(x', y') > 0$  for all  $n \geq 1$ . We may now repeat the argument presented in the previous paragraph to construct a path in the set  $\mathcal{V}_j^{(p-1)}$  from  $x$  to  $x'$ , and a second one in the set  $\mathcal{V}_k^{(p-1)}$  from  $y'$  to  $y$ . Chaining the paths yields a path  $(x = z_0, z_1, \dots, z_\ell = y)$  such that  $z_i \in \mathcal{V}_m^{(p)}$ ,  $R_n^{(p)}(z_i, z_{i+1}) > 0$  for all  $0 \leq i < \ell$  and  $n \geq 1$ . This completes the proof of the lemma.  $\square$

**Lemma 3.2.** *Fix  $1 \leq p \leq q$ . Then, for all  $x, y \in \mathcal{V}^{(p)}$ ,*

$$\lim_{n \rightarrow \infty} R_n^{(p)}(x, y) = \mathbb{R}_0(x, y).$$

*Proof.* Fix  $1 \leq p \leq q$  and  $x, y \in \mathcal{V}^{(p)}$ . By (3.1), decomposing the probability appearing on the right-hand side of this equation according to the first jump yields that

$$R_n^{(p)}(x, y) = R_n(x, y) + \sum_{z \neq y} R_n(x, z) \mathbf{P}_z^n [H_y = H_{\mathcal{V}^{(p)}}].$$

The first term converges to  $\mathbb{R}_0(x, y)$ . As  $x \in \mathcal{V}^{(p)}$  and (by the tree construction)  $\mathcal{V}^{(p)}$  is the union of some sets  $\mathcal{V}_k$ ,  $k \in S_1$ ,  $x \in \mathcal{V}_j$  for some  $j \in S_1$ . The probability on the second term vanishes if  $z \in \mathcal{V}^{(p)}$ . We may therefore restrict the sum to  $z \notin \mathcal{V}^{(p)}$ , or to  $\mathcal{V}_j^c$  (because  $\mathcal{V}_j \subset \mathcal{V}^{(p)}$ ). However, by definition of  $\mathcal{V}_j$ ,  $R_n(x, z) \rightarrow 0$  for all  $z \notin \mathcal{V}_j$ . Thus, the second term of the previous displayed formula vanishes, which completes the proof of the lemma.  $\square$

**Lemma 3.3.** *Fix  $1 \leq p \leq q$ . Then,*

$$\limsup_{n \rightarrow \infty} \theta_n^{(p)} R_n^{(p)}(x, y) < \infty$$

for all  $k \neq j \in S_p$ ,  $x \in \mathcal{V}_j^{(p)}$ ,  $y \in \mathcal{V}_k^{(p)}$ .

*Proof.* As  $y \in \mathcal{V}_k^{(p)}$ ,  $R_n^{(p)}(x, y) \leq R_n^{(p)}(x, \mathcal{V}_k^{(p)})$ . By (2.13), there exists a finite constant  $C_0$  such that

$$\theta_n^{(p)} R_n^{(p)}(x, y) \leq C_0 \theta_n^{(p)} \sum_{z \in \mathcal{V}_j^{(p)}} \frac{\pi_n(z)}{\pi_n(\mathcal{V}_j^{(p)})} R_n^{(p)}(z, \mathcal{V}_k^{(p)}).$$

It remains to recall (3.3) to complete the proof.  $\square$

**Lemma 3.4.** *For all  $1 \leq p \leq q$ ,  $k \neq j \in S_1$  such that  $\mathcal{V}_j^{(1)} \cup \mathcal{V}_k^{(1)} \subset \mathcal{V}^{(p)}$ ,  $x \in \mathcal{V}_j^{(1)}$ , the sequence  $\theta_n^{(1)} R_n^{(p)}(x, \mathcal{V}_k^{(1)})$  is bounded.*

*Proof.* The proof is by induction on  $p$ . For  $p = 1$ , the assertion of the lemma follows from (3.3), (3.2) and (2.13).

Fix  $p > 1$ , and assume that the assertion of the lemma holds for  $1 \leq q < p$ . Fix  $k \neq j \in S_1$  such that  $\mathcal{V}_j^{(1)} \cup \mathcal{V}_k^{(1)} \subset \mathcal{V}^{(p)}$ ,  $x \in \mathcal{V}_j^{(1)}$ . By (3.1),

$$R_n^{(p)}(x, \mathcal{V}_k^{(1)}) = \lambda_n(x) \mathbf{P}_x^n [H_{\mathcal{V}_k^{(1)}} = H_{\mathcal{V}^{(p)}}^+].$$

Recall that  $\mathcal{V}^{(p)} \subset \mathcal{V}^{(p-1)}$ . Assume that  $H_{\mathcal{V}^{(p)}}^+ = H_{\mathcal{V}^{(p-1)}}^+$ . Later we consider the case  $H_{\mathcal{V}^{(p)}}^+ > H_{\mathcal{V}^{(p-1)}}^+$ . In the first case, we need to estimate

$$\begin{aligned} & \lambda_n(x) \mathbf{P}_x^n [H_{\mathcal{V}_k^{(1)}} = H_{\mathcal{V}^{(p)}}^+, H_{\mathcal{V}^{(p)}}^+ = H_{\mathcal{V}^{(p-1)}}^+] \\ & \leq \lambda_n(x) \mathbf{P}_x^n [H_{\mathcal{V}_k^{(1)}} = H_{\mathcal{V}^{(p-1)}}^+] = R_n^{(p-1)}(x, \mathcal{V}_k^{(1)}). \end{aligned}$$

By the induction hypothesis this later quantity multiplied by  $\theta_n^{(1)}$  is bounded.

It remains to estimate the expression

$$\lambda_n(x) \mathbf{P}_x^n [H_{\mathcal{V}^{(p-1)}}^+ < H_{\mathcal{V}^{(p)}}^+].$$

By construction, there exists  $S'_{p-1} \subset S_{p-1}$  such that  $\mathcal{V}^{(p-1)} \setminus \mathcal{V}^{(p)} = \cup_{\ell \in S'_{p-1}} \mathcal{V}_\ell^{(p-1)}$ . Mind that  $S'_{p-1}$  consists of the transient points of the  $S_{p-1}$ -valued Markov chain  $\mathbb{X}_t^{(p-1)}$ . Since  $x \in \mathcal{V}^{(p)} \subset \mathcal{V}^{(p-1)}$ , let  $m \in S_{p-1} \setminus S'_{p-1}$  such that  $x \in \mathcal{V}_m^{(p-1)}$ . With this notation, the previous term is bounded by

$$\sum_{\ell \in S'_{p-1}} \lambda_n(x) \mathbf{P}_x^n [H_{\mathcal{V}_\ell^{(p-1)}} = H_{\mathcal{V}^{(p-1)}}^+] = \sum_{\ell \in S'_{p-1}} R_n^{(p-1)}(x, \mathcal{V}_\ell^{(p-1)}).$$

By (3.3), (3.2) and (2.13), the limit as  $n \rightarrow \infty$  of the previous expression multiplied by  $\theta_n^{(p-1)}$  is bounded by

$$\sum_{\ell \in S'_{p-1}} r^{(p-1)}(m, \ell).$$

This sum vanishes because  $m$  is a recurrent point of the chains  $\mathbb{X}_t^{(p-1)}$  and  $S'_{p-1}$  is a transient subset. To complete the proof of the lemma, it remains to recall that  $\theta_n^{(1)} \leq \theta_n^{(p-1)}$ .  $\square$

**Lemma 3.5.** *Fix  $1 \leq p < q \leq q$ . Then,  $r^{(p)}(i, k) = 0$  for all  $k \neq i \in S_p$  such that  $\mathcal{V}_k^{(p)} \not\subset \mathcal{V}^{(q)}$ ,  $\mathcal{V}_i^{(p)} \subset \mathcal{V}^{(q)}$ .*

*Proof.* As  $\mathcal{V}_k^{(p)} \subset \mathcal{V}^{(p)}$  and  $\mathcal{V}_k^{(p)} \not\subset \mathcal{V}^{(q)}$ , by the tree construction there exists  $p \leq p' < q$  such that  $\mathcal{V}_k^{(p)} \subset \mathcal{V}^{(p')}$  and  $\mathcal{V}_k^{(p)} \not\subset \mathcal{V}^{(p'+1)}$ . Since  $\mathcal{V}_k^{(p)} \subset \mathcal{V}^{(p')}$ , there exists

$\ell \in S_{p'}$  such that  $\mathcal{V}_k^{(p)} \subset \mathcal{V}_\ell^{(p')}$ . As  $\mathcal{V}_k^{(p)} \not\subset \mathcal{V}^{(p'+1)}$ ,  $\ell$  is a transient state for the Markov chain  $\mathbb{X}_t^{(p')}$ .

On the other hand, as  $\mathcal{V}_i^{(p)} \subset \mathcal{V}^{(q)}$  and  $\mathcal{V}^{(q)} \subset \mathcal{V}^{(p')}$ ,  $\mathcal{V}_i^{(p)} \subset \mathcal{V}^{(p')}$ . Thus, there exists  $m \in S_{p'}$  such that  $\mathcal{V}_i^{(p)} \subset \mathcal{V}_m^{(p')}$ . As  $\mathcal{V}_i^{(p)} \subset \mathcal{V}^{(q)} \subset \mathcal{V}^{(p'+1)}$ ,  $m$  is a recurrent state for the Markov chain  $\mathbb{X}_t^{(p')}$ . In particular,  $m \neq \ell$ .

As  $m$  is recurrent and  $\ell$  transient for the Markov chain  $\mathbb{X}_t^{(p')}$ ,  $r^{(p')}(m, \ell) = 0$ . Thus, by (3.3) and (3.2),

$$0 = \lim_{n \rightarrow \infty} \theta_n^{(p')} r_n^{(p')}(m, \ell) = \lim_{n \rightarrow \infty} \theta_n^{(p')} \frac{1}{\pi_n(\mathcal{V}_m^{(p')})} \sum_{x \in \mathcal{V}_m^{(p')}} \pi_n(x) R_n^{(p')}(x, \mathcal{V}_\ell^{(p')}).$$

By (3.1), for  $n$  fixed the expression on right-hand side is equal to

$$\theta_n^{(p')} \frac{1}{\pi_n(\mathcal{V}_m^{(p')})} \sum_{x \in \mathcal{V}_m^{(p')}} \pi_n(x) \lambda_n(x) \mathbf{P}_x^n [H_{\mathcal{V}_\ell^{(p')}} = H_{\mathcal{V}^{(p')}}^+]. \quad (3.4)$$

Since  $\mathcal{V}_k^{(p)} \subset \mathcal{V}_\ell^{(p')}$  and  $\mathcal{V}^{(p)} \supset \mathcal{V}^{(p')}$ ,

$$\mathbf{P}_x^n [H_{\mathcal{V}_k^{(p)}} = H_{\mathcal{V}^{(p)}}^+] \leq \mathbf{P}_x^n [H_{\mathcal{V}_\ell^{(p')}} = H_{\mathcal{V}^{(p')}}^+].$$

Hence, as  $\theta_n^{(p')} \geq \theta_n^{(p)}$  and  $\mathcal{V}_m^{(p')} \supset \mathcal{V}_i^{(p)}$ , by (2.13), (3.4) is bounded below by

$$c_0 \theta_n^{(p)} \sum_{x \in \mathcal{V}_i^{(p)}} \lambda_n(x) \mathbf{P}_x^n [H_{\mathcal{V}_k^{(p)}} = H_{\mathcal{V}^{(p)}}^+]$$

for some positive constant  $c_0$ . This expression is clearly bounded below by

$$c_0 \theta_n^{(p)} \frac{1}{\pi_n(\mathcal{V}_i^{(p)})} \sum_{x \in \mathcal{V}_i^{(p)}} \pi_n(x) \lambda_n(x) \mathbf{P}_x^n [H_{\mathcal{V}_k^{(p)}} = H_{\mathcal{V}^{(p)}}^+] = c_0 r_n^{(p)}(i, k).$$

Collecting the previous estimates yields that this expression vanishes as  $n \rightarrow \infty$ , as claimed.  $\square$

For  $1 \leq p < q \leq \mathfrak{q}$  and  $i \neq j \in S_p$ . Assume that  $\mathcal{V}_i^{(p)}$  and  $\mathcal{V}_j^{(p)}$  are contained in  $\mathcal{V}^{(q)}$ :  $\mathcal{V}_i^{(p)} \cup \mathcal{V}_j^{(p)} \subset \mathcal{V}^{(q)}$ . Let

$$r_n^{p,q}(i, j) := \frac{1}{\pi_n(\mathcal{V}_i^{(p)})} \sum_{x \in \mathcal{V}_i^{(p)}} \pi_n(x) \sum_{y \in \mathcal{V}_j^{(p)}} R_n^{(q)}(x, y).$$

The difference with respect to  $r_n^{(p)}(i, j)$  is that we replaced  $R_n^{(p)}(x, y)$  by  $R_n^{(q)}(x, y)$ , that is, the trace on  $\mathcal{V}^{(p)}$  by the one on the smaller set  $\mathcal{V}^{(q)}$ .

**Corollary 3.6.** *Fix  $1 \leq p < q \leq \mathfrak{q}$ ,  $i \neq j \in S_p$ . Assume that there exists  $m \in S_q$  such that  $\mathcal{V}_i^{(p)} \cup \mathcal{V}_j^{(p)} \subset \mathcal{V}_m^{(q)}$ . Then,*

$$\lim_{n \rightarrow \infty} \theta_n^{(p)} r_n^{p,q}(i, j) = r^{(p)}(i, j).$$

*Proof.* By (3.3), (3.2) and (3.1),

$$r^p(i, j) = \lim_{n \rightarrow \infty} \frac{\theta_n^{(p)}}{\pi_n(\mathcal{V}_i^{(p)})} \sum_{x \in \mathcal{V}_i^{(p)}} \pi_n(x) \lambda_n(x) \mathbf{P}_x^n [H_{\mathcal{V}_j^{(p)}} = H_{\mathcal{V}^{(p)}}^+]. \quad (3.5)$$

Let  $S_{p,q} := \{k \in S_p : \mathcal{V}_k^{(p)} \subset \mathcal{V}^{(q)}\}$ ,  $\mathcal{V}^{(p,q)} := \cup_{k \in S_{p,q}} \mathcal{V}_k^{(p)}$ ,  $\mathcal{U}^{(p,q)} := \mathcal{V}^{(p)} \setminus \mathcal{V}^{(q)}$ . By Lemma 3.5,

$$\lim_{n \rightarrow \infty} \frac{\theta_n^{(p)}}{\pi_n(\mathcal{V}_i^{(p)})} \sum_{x \in \mathcal{V}_i^{(p)}} \pi_n(x) \lambda_n(x) \mathbf{P}_x^n [H_{\mathcal{U}^{(p,q)}} = H_{\mathcal{V}^{(p)}}^+] = 0. \quad (3.6)$$

Since  $\mathcal{V}^{(p)} = \mathcal{V}^{(p,q)} \cup \mathcal{U}^{(p,q)}$ ,  $\mathcal{V}^{(p,q)} \cap \mathcal{U}^{(p,q)} = \emptyset$ , the sets  $\{H_{\mathcal{U}^{(p,q)}} = H_{\mathcal{V}^{(p)}}^+\}$  and  $\{H_{\mathcal{V}^{(q)}}^+ = H_{\mathcal{V}^{(p)}}^+\}$  form a partition of the space. Decomposing the event appearing in (3.5) according to this partition, by (3.6),

$$r^p(i, j) = \lim_{n \rightarrow \infty} \frac{\theta_n^{(p)}}{\pi_n(\mathcal{V}_i^{(p)})} \sum_{x \in \mathcal{V}_i^{(p)}} \pi_n(x) \lambda_n(x) \mathbf{P}_x^n [H_{\mathcal{V}_j^{(p)}} = H_{\mathcal{V}^{(q)}}^+, H_{\mathcal{V}^{(q)}}^+ = H_{\mathcal{V}^{(p)}}^+].$$

By (3.6) once more,

$$r^p(i, j) = \lim_{n \rightarrow \infty} \frac{\theta_n^{(p)}}{\pi_n(\mathcal{V}_i^{(p)})} \sum_{x \in \mathcal{V}_i^{(p)}} \pi_n(x) \lambda_n(x) \mathbf{P}_x^n [H_{\mathcal{V}_j^{(p)}} = H_{\mathcal{V}^{(q)}}^+].$$

By definition, the right-hand side is  $\lim_{n \rightarrow \infty} \theta_n^{(p)} r_n^{p,q}(i, j)$ , which completes the proof of the lemma.  $\square$

**Lemma 3.7.** *Fix  $1 \leq p < q \leq \mathfrak{q}$ . Then,*

$$r^{(p)}(i, k) = 0$$

for all  $k \neq i \in S_p$  such that  $\mathcal{V}_i^{(p)} \subset \mathcal{V}_a^{(q)}$ ,  $\mathcal{V}_k^{(p)} \subset \mathcal{V}_b^{(q)}$  for some  $a \neq b \in S_q$ .

*Proof.* Fix  $1 \leq p < q \leq \mathfrak{q}$ , and  $k \neq i \in S_p$  such that  $\mathcal{V}_i^{(p)} \subset \mathcal{V}_a^{(q)}$ ,  $\mathcal{V}_k^{(p)} \subset \mathcal{V}_b^{(q)}$  for some  $a \neq b \in S_q$ . Both states  $i$  and  $k$  are recurrent for the chain  $\mathbb{X}_t^{(p)}$  because if one of them was transient it would not belong to  $\mathcal{V}^{(p+1)} \supset \mathcal{V}^{(q)}$ .

Suppose by contradiction that  $r^{(p)}(i, k) > 0$ . Hence, since both states are recurrent, they belong to the same irreducible class. In particular, there exists  $m \in S_{p+1}$  such that  $\mathcal{V}_i^{(p)} \cup \mathcal{V}_k^{(p)} \subset \mathcal{V}_m^{(p+1)}$ , so that  $\mathcal{V}_i^{(p)} \cup \mathcal{V}_k^{(p)} \subset \mathcal{V}_c^{(q)}$  for some  $c \in S_q$ , in contradictions with the hypotheses.  $\square$

#### 4. $\Gamma$ – lim sup OF THE TRACE

The main result of this section, Proposition 4.1, states that  $\mathcal{J}^{(p)}$  is a  $\Gamma$  – lim sup for the sequence  $\theta_n^{(p)} \mathcal{J}_n^{(p)}$ . Here,  $\mathcal{J}_n^{(p)}$  stands for the large deviations rate functionals of the trace processes  $Y_t^{n,p}$ . We assume below that the reader is familiar with the notation and results presented in the appendix.

Fix  $1 \leq p \leq \mathfrak{q}$ . Denote by  $\mathcal{J}_n^{(p)} : \mathcal{P}(\mathcal{V}^{(p)}) \rightarrow [0, +\infty)$  the occupation time large deviations rate functional of the trace process  $Y_t^{n,p}$ :

$$\mathcal{J}_n^{(p)}(\mu) := \sup_H - \sum_{x \in \mathcal{V}^{(p)}} \mu(x) e^{-H(x)} [(\mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n) e^H](x), \quad (4.1)$$

where the supremum is carried over all functions  $H : \mathcal{V}^{(p)} \rightarrow \mathbb{R}$  and  $\mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n$ , introduced and examined in Appendix C, is the generator of the trace process  $Y_t^{n,p}$ . The main result of this section reads as follows.

**Proposition 4.1.** *For all  $\mu \in \mathcal{P}(\mathcal{V}^{(p)})$ , there exists a sequence of measures  $\mu_n \in \mathcal{P}(\mathcal{V}^{(p)})$  such that  $\mu_n \rightarrow \mu$  and*

$$\limsup_{n \rightarrow \infty} \theta_n^{(p)} \mathcal{J}_n^{(p)}(\mu_n) \leq \mathcal{J}^{(p)}(\mu) .$$

The proof of Proposition 4.1 is divided in several lemmata. Fix  $1 \leq p \leq \mathfrak{q}$  and a measure  $\mu \in \mathcal{P}(\mathcal{V}^{(p)})$  which can be represented as  $\mu = \sum_{j \in S_p} \omega_j \pi_j^{(p)}$  for some  $\omega \in \mathcal{P}(S_p)$  such that  $\omega_j > 0$  for all  $j \in S_p$ . We first construct a sequence  $\mu_n \in \mathcal{P}(\mathcal{V}^{(p)})$  which converges to  $\mu$ .

Denote by  $\mathfrak{D}_a^{(p)}$ ,  $1 \leq a \leq \ell_p$ , the equivalent classes of the Markov chain  $\mathbb{X}_t^{(p)}$ , by  $\mathfrak{D}_a^{(p)}$ ,  $1 \leq a \leq m_p$ , the ones which are not singletons, and by  $S_p^{\text{sgl}}$  the set of states  $j \in S_p$  such that  $\{j\}$  is an equivalent class. Clearly,

$$S_p = \bigcup_{a=1}^{\ell_p} \mathfrak{D}_a^{(p)} = S_p^{\text{sgl}} \cup \bigcup_{a=1}^{m_p} \mathfrak{D}_a^{(p)} . \quad (4.2)$$

Recall from Appendix A the definition of a Markov chain reflected at a set and the notation used to represent its generator. Denote by  $\mathbb{L}_a^{(p)}$ ,  $1 \leq a \leq m_p$ , the generator  $\mathbb{L}^{(p)}$  reflected at  $\mathfrak{D}_a^{(p)}$ :  $\mathbb{L}_a^{(p)} := \mathfrak{R}_{\mathfrak{D}_a^{(p)}} \mathbb{L}^{(p)}$ . As  $\omega_j > 0$  for all  $j \in \mathfrak{D}_a^{(p)}$ , by Lemma A.3, there exists  $\mathbf{h}_a: \mathfrak{D}_a^{(p)} \rightarrow \mathbb{R}$  which solves the optimal problem (A.3) for  $I_{\mathbb{L}_a^{(p)}}(\omega)$ .

Let  $\mathcal{W}_a^{(p)} := \bigcup_{j \in \mathfrak{D}_a^{(p)}} \mathcal{V}_j^{(p)}$ ,  $1 \leq a \leq m_p$ . The generator of the trace process  $Y_t^{n,p}$  reflected at the set  $\mathcal{W}_a^{(p)}$  is denoted by  $\mathfrak{R}_{\mathcal{W}_a^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n$ .

**Lemma 4.2.** *The Markov chain associated to the generator  $\mathfrak{R}_{\mathcal{W}_a^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n$  is irreducible.*

*Proof.* Recall from [2, Proposition 6.1] that the Markov chain induced by the trace generator  $\mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n$  is irreducible. Since  $\mathfrak{D}_a^{(p)}$  is an equivalent class for the Markov chain  $\mathbb{X}_t^{(p)}$ , the argument presented in the proof of Lemma 3.1 yields that the Markov chain induced by the reflected generator  $\mathfrak{R}_{\mathcal{W}_a^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n$  is also irreducible.  $\square$

Denote by  $H_a: \mathcal{W}_a^{(p)} \rightarrow \mathbb{R}$  the function given by

$$H_a = \sum_{j \in \mathfrak{D}_a^{(p)}} \mathbf{h}_a(j) \chi_{\mathcal{V}_j^{(p)}} , \quad (4.3)$$

where, recall,  $\chi_{\mathcal{A}}$  stands for the indicator function of the set  $\mathcal{A}$ . Recall from (A.5) the definition of a tilted generator  $\mathfrak{M}_G L$ , and consider the generator  $\mathfrak{M}_{H_a} \mathfrak{R}_{\mathcal{W}_a^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n$ . Since tilting the generator does not affect its irreducibility, it follows from the previous result that the Markov chain associated to the generator  $\mathfrak{M}_{H_a} \mathfrak{R}_{\mathcal{W}_a^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n$  is also irreducible. Denote by  $\mu_n^{p,a} \in \mathcal{P}(\mathcal{W}_a^{(p)})$  its stationary state.

**Lemma 4.3.** *The sequence of probability measures  $\mu_n^{p,a}$  converges to  $\sum_{j \in \mathfrak{D}_a^{(p)}} \omega_j^{(a)} \pi_j^{(p)}$ , where  $\omega_j^{(a)} = \omega_j / \Omega_a$ ,  $\Omega_a = \sum_{j \in \mathfrak{D}_a^{(p)}} \omega_j$ .*

*Proof.* Since  $\mathcal{P}(\mathcal{W}_a^{(p)})$  is compact for the weak topology, it is enough to prove uniqueness of limit points. consider a subsequence of  $\mu_n^{p,a}$ , still denoted by  $\mu_n^{p,a}$ , which converges to a limit denoted by  $\nu \in \mathcal{P}(\mathcal{W}_a^{(p)})$ .



**Step 1:  $\nu$  on the sets  $\mathcal{V}_i^{(1)}$ .** The set  $\mathcal{W}_a^{(p)}$  is the union of sets  $\mathcal{V}_i^{(p)}$ , which in turn are formed by sets  $\mathcal{V}_j^{(1)}$ . Hence,  $\mathcal{W}_a^{(p)} = \cup_{i \in S_1^a} \mathcal{V}_i^{(1)}$  for some subset  $S_1^a$  of  $S_1$ .

Fix  $x \in \mathcal{W}_a^{(p)}$  and assume that  $x \in \mathcal{V}_k^{(1)}$  for some  $k \in S_1^a$ . Since  $\mu_n^{p,a}$  is the stationary state for the chain induced by the generator  $\mathfrak{M}_{H_a} \mathfrak{R}_{\mathcal{W}_a^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n$ ,

$$\sum_{y \in \mathcal{W}_a^{(p)}} \mu_n^{p,a}(y) R_n^{(p)}(y, x) e^{H_a(x) - H_a(y)} = \sum_{y \in \mathcal{W}_a^{(p)}} \mu_n^{p,a}(x) R_n^{(p)}(x, y) e^{H_a(y) - H_a(x)}.$$

Since  $\mu_n^{p,a} \rightarrow \nu$ , taking  $n \rightarrow \infty$  in the previous formula, by Lemma 3.2,

$$\sum_{y \in \mathcal{W}_a^{(p)}} \nu(y) \mathbb{R}_0(y, x) e^{H_a(x) - H_a(y)} = \sum_{y \in \mathcal{W}_a^{(p)}} \nu(x) \mathbb{R}_0(x, y) e^{H_a(y) - H_a(x)}.$$

By definition of the sets  $\mathcal{V}_i^{(1)}$ ,  $\mathbb{R}_0(z, w) = 0$  if  $z \in \mathcal{V}_i^{(1)}$ ,  $w \in \mathcal{V}_{i'}^{(1)}$  and  $i \neq i'$ . Hence, as  $x \in \mathcal{V}_k^{(1)}$  and  $H_a$  is constant on each set  $\mathcal{V}_j^{(1)}$ , the previous identity becomes

$$\sum_{y \in \mathcal{V}_k^{(1)}} \nu(y) \mathbb{R}_0(y, x) = \sum_{y \in \mathcal{V}_k^{(1)}} \nu(x) \mathbb{R}_0(x, y).$$

Therefore, the measure  $\nu$  restricted to  $\mathcal{V}_k^{(1)}$  is a stationary measure for the Markov chain  $\mathbb{X}_t$  restricted to  $\mathcal{V}_k^{(1)}$ . Since this process is irreducible on  $\mathcal{V}_k^{(1)}$ , by the definition of  $\pi_k^{(1)}$  given right after (2.11),  $\nu(\cdot) = \nu(\mathcal{V}_k^{(1)}) \pi_k^{(1)}(\cdot)$ . Hence,  $\nu$  is a convex combination of the stationary states  $\pi_\ell^{(1)}$ :

$$\nu = \sum_{k \in S_1^a} \vartheta_1(k) \pi_k^{(1)} \quad (4.4)$$

for some probability measure  $\vartheta_1$  on  $S_1^a$ .

**Step 2: An equation for  $\vartheta_1$ .** Fix  $1 \leq r \leq p$ , and let  $S_r^a := \{i \in S_r : \mathcal{V}_i^{(r)} \subset \mathcal{W}_a^{(p)}\}$  so that  $\mathcal{W}_a^{(p)} = \cup_{i \in S_r^a} \mathcal{V}_i^{(r)}$ . Fix  $\mathbf{g} : S_r^a \rightarrow \mathbb{R}$  and let  $G : \mathcal{W}_a^{(p)} \rightarrow \mathbb{R}$  be given by  $G = \sum_{j \in S_r^a} \mathbf{g}(j) \chi_{\mathcal{V}_j^{(r)}}$ . As  $\mu_n^{p,a}$  is a stationary state,

$$\sum_{x \in \mathcal{W}_a^{(p)}} \mu_n^{p,a}(x) \sum_{y \in \mathcal{W}_a^{(p)}} R_n^{(p)}(x, y) e^{H_a(y) - H_a(x)} [G(y) - G(x)] = 0.$$

Since  $H_a$  is constant on the sets  $\mathcal{V}_\ell^{(p)}$ , it is also constant on the sets  $\mathcal{V}_j^{(r)}$ , which are subsets of the former sets. By definition of  $G$  and  $H_a$ , this identity can be written as

$$\sum_{j \in S_r^a} \sum_{k \in S_r^a \setminus \{j\}} e^{\mathbf{f}_a(k) - \mathbf{f}_a(j)} [\mathbf{g}(k) - \mathbf{g}(j)] \sum_{x \in \mathcal{V}_j^{(r)}} \mu_n^{p,a}(x) R_n^{(p)}(x, \mathcal{V}_k^{(r)}) = 0. \quad (4.5)$$

Here  $\mathbf{f}_a : S_r^a \rightarrow \mathbb{R}$  is the function defined by  $\mathbf{f}_a(j) = \mathbf{h}_a(m)$  for all  $j \in S_r^{a,m}$ , where  $S_r^{a,m} := \{i \in S_r : \mathcal{V}_i^{(r)} \subset \mathcal{V}_m^{(p)}\}$ ,  $m \in S_p^a$ .

If  $p = 1$ , jump to Step 5. Assume below that  $p > 1$  and set  $r = 1$ . We claim that for each  $j \neq k \in S_1^a$ ,

$$\lim_{n \rightarrow \infty} \theta_n^{(1)} \sum_{x \in \mathcal{V}_j^{(1)}} \mu_n^{p,a}(x) R_n^{(p)}(x, \mathcal{V}_k^{(1)}) = \vartheta_1(j) r^{(1)}(j, k), \quad (4.6)$$

where  $\vartheta_1 \in \mathcal{P}(S_1^a)$  is the probability measure obtained in the first step.

To prove (4.6), rewrite the expression on the left-hand side as

$$\theta_n^{(1)} \sum_{x \in \mathcal{V}_j^{(1)}} \frac{\pi_n(\mathcal{V}_j^{(1)})}{\pi_n(x)} \mu_n^{p,a}(x) \frac{\pi_n(x)}{\pi_n(\mathcal{V}_j^{(1)})} R_n^{(p)}(x, \mathcal{V}_k^{(1)}).$$

By the first part of the proof,  $\mu_n^{p,a}(x) \rightarrow \vartheta_1(j) \pi_j^{(1)}(x)$ . By (2.13),  $\pi_n(\mathcal{V}_j^{(1)})/\pi_n(x) \rightarrow 1/\pi_j^{(1)}(x)$ . By Lemma 3.4, the sequence  $\theta_n^{(1)} R_n^{(p)}(x, \mathcal{V}_k^{(1)})$  is bounded. Therefore, by Corollary 3.6, (4.6) holds.

By (4.5) and (4.6),

$$\sum_{j \in S_1^a} \sum_{k \in S_1^a \setminus \{j\}} \vartheta_1(j) r^{(1)}(j, k) e^{\mathbf{f}_a(k) - \mathbf{f}_a(j)} [\mathbf{g}(k) - \mathbf{g}(j)] = 0$$

for all functions  $\mathbf{g}: S_1^a \rightarrow \mathbb{R}$ .

By Lemma 3.7,  $r^{(1)}(j, k) = 0$  for all  $j \in S_1^{a,m}$ ,  $k \in S_1^{a,\ell}$  and  $m \neq \ell$ . We may therefore rewrite the previous identity as

$$\sum_{m \in S_p^a} \sum_{j \in S_1^{a,m}} \sum_{k \in S_1^{a,m} \setminus \{j\}} \vartheta_1(j) r^{(1)}(j, k) e^{\mathbf{f}_a(k) - \mathbf{f}_a(j)} [\mathbf{g}(k) - \mathbf{g}(j)] = 0.$$

As the function  $\mathbf{f}_a$  is constant on each  $S_1^{a,m}$ , we may remove the exponential from the previous equation and rewrite the identity as

$$\sum_{j \in S_1^a} \sum_{k \in S_1^a \setminus \{j\}} \vartheta_1(j) r^{(1)}(j, k) [\mathbf{g}(k) - \mathbf{g}(j)] = 0$$

for all functions  $\mathbf{g}: S_1^a \rightarrow \mathbb{R}$ . Hence,  $\vartheta_1(\cdot)$  is a stationary state for the chain  $\mathbb{X}_t^{(1)}$  reflected at  $S_1^a$ .

**Step 3: from  $\vartheta_1$  to  $\vartheta_2$**  Recall that if  $p = 1$ , the proof continues at Step 5. By its definition, given a few lines above equation (2.12),  $M_m^{(1)}$ ,  $m \in S_2$ , represents the stationary state of the chain  $\mathbb{X}_t^{(1)}$  reflected at  $S_1^{2,m} = \{j \in S_1 : \mathcal{V}_j^{(1)} \subset \mathcal{V}_m^{(2)}\}$ . Since  $\vartheta_1$  is a stationary state of the chain  $\mathbb{X}_t^{(1)}$  whose support is contained in  $S_1^a$ ,  $\vartheta_1$  is a convex combination of the measures  $M_m^{(1)}$ ,  $m \in S_2^a$ :

$$\vartheta_1(\cdot) = \sum_{m \in S_2^a} \vartheta_2(m) M_m^{(1)}(\cdot)$$

for a probability measure  $\vartheta_2 \in \mathcal{P}(S_2^a)$ . Thus,

$$\nu(\cdot) = \sum_{k \in S_1^a} \vartheta_1(k) \pi_k^{(1)}(\cdot) = \sum_{k \in S_1^a} \sum_{m \in S_2^a} \vartheta_2(m) M_m^{(1)}(k) \pi_k^{(1)}(\cdot).$$

Changing the order of summation, and recalling the definition of the measures  $\pi_m^{(2)}$  yields that

$$\nu(\cdot) = \sum_{m \in S_2^a} \vartheta_2(m) \sum_{k \in S_1^{2,m}} M_m^{(1)}(k) \pi_k^{(1)}(\cdot) = \sum_{m \in S_2^a} \vartheta_2(m) \pi_m^{(2)}(\cdot). \quad (4.7)$$

Step 2 and 3 permitted to pass from (4.4) to (4.7). That is, at the end of Step 1 we obtained that  $\nu$  is a convex combination of the measures  $\pi_j^{(1)}$ . We now have shown that it is, actually, a convex combination of the measures  $\pi_j^{(2)}$ . Next step consist in iterating this argument to obtain that  $\nu$  is, actually, a convex combination of the measures  $\pi_j^{(p)}$ .

**Step 4: An iteration.** If  $p = 2$  the proof continues at Step 5. Assume here that  $p > 2$ , and suppose that we proved that

$$\nu(\cdot) = \sum_{j \in S_s^a} \vartheta_s(j) \pi_j^{(s)}(\cdot)$$

for some  $2 \leq s < p$  and some probability measure  $\vartheta_s \in \mathcal{P}(S_s^a)$ .

Resume the proof of Step 2 at equation (4.5) with  $r = s$ . Recall the argument presented to derive (4.6). Fix  $j \neq k \in S_s^a$ . Inserting  $\pi_n(\mathcal{V}_j^{(s)})/\pi_n(x)$  instead of  $\pi_n(\mathcal{V}_j^{(1)})/\pi_n(x)$  yields that

$$\lim_{n \rightarrow \infty} \theta_n^{(s)} \sum_{x \in \mathcal{V}_j^{(s)}} \mu_n^{p,a}(x) R_n^{(p)}(x, \mathcal{V}_k^{(s)}) = \vartheta_s(j) r^{(s)}(j, k). \quad (4.8)$$

By Lemma 3.7,  $r^{(s)}(j, k) = 0$  if  $j \in S_s^{a,\ell}$ ,  $k \in S_s^{a,m}$  and  $\ell \neq m$ . Hence, applying the arguments presented in Step 3, one gets that

$$\nu(\cdot) = \sum_{j \in S_{s+1}^a} \vartheta_{s+1}(j) \pi_j^{(s+1)}(\cdot)$$

for some some probability measure  $\vartheta_{s+1} \in \mathcal{P}(S_{s+1}^a)$ .

Iterating this procedure yields that

$$\nu(\cdot) = \sum_{j \in S_p^a} \vartheta_p(j) \pi_j^{(p)}(\cdot) \quad (4.9)$$

for some some probability measure  $\vartheta_p \in \mathcal{P}(S_p^a)$ .

**Step 5: Conclusion.** Applying again the arguments presented in Step 2 yields that

$$\sum_{j \in S_p^a} \sum_{k \in S_p^a \setminus \{j\}} \vartheta_p(j) r^{(p)}(j, k) e^{\mathbf{h}_a(k) - \mathbf{h}_a(j)} [\mathbf{g}(k) - \mathbf{g}(j)] = 0$$

for all function  $\mathbf{g}: S_p^a \rightarrow \mathbb{R}$ . In particular,  $\vartheta_p$  is a stationary state for the  $S_p^a$ -valued Markov chain which jumps from  $j$  to  $k$  at rate  $r_{\mathbf{h}_a}^{(p)}(j, k) := r^{(p)}(j, k) \exp\{\mathbf{h}_a(k) - \mathbf{h}_a(j)\}$ . Since the chain  $\mathbb{X}_t^{(p)}$  is irreducible on  $S_p^a$ , so is the one with jump rates  $r_{\mathbf{h}_a}^{(p)}(j, k)$ . Hence,  $\vartheta_p$  is the unique stationary state. Since  $\omega$  is also stationary,  $\vartheta_p(\cdot) = \Omega_a^{-1} \omega(\cdot)$ .

This proves the uniqueness of limit points of the sequence  $\mu_n^{p,a}$  and completes the proof of the lemma.  $\square$

Fix  $j \in S_p^{\text{sgl}}$ . By Lemma 3.1, the Markov chain induced by the generator  $\mathfrak{R}_{\mathcal{V}_j^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n$  is irreducible. Denote by  $\nu_n^{p,j} \in \mathcal{P}(\mathcal{V}_j^{(p)})$  its stationary state.

**Lemma 4.4.** *For each  $j \in S_p^{\text{sgl}}$ , the sequence of probability measures  $\nu_n^{p,j}$  converges to  $\pi_j^{(p)}$ .*

*Proof.* The proof is similar (and simpler) than the one of Lemma 4.3. It amounts to take  $H = 0$  and  $\mathcal{W}_a^{(p)} = \mathcal{V}_j^{(p)}$ .  $\square$

Consider the measures  $\mu_n^{p,a}$ ,  $\nu_n^{p,j}$ ,  $1 \leq a \leq m_p$ ,  $j \in S_p^{\text{sgl}}$ , as probability measures on  $\mathcal{V}^{(p)}$ , and let  $\mu_n \in \mathcal{P}(\mathcal{V}^{(p)})$  be the measure given by

$$\mu_n(\cdot) = \sum_{a=1}^{m_p} \Omega_a \mu_n^{p,a}(\cdot) + \sum_{j \in S_p^{\text{sgl}}} \omega_j \nu_n^{p,j}(\cdot). \quad (4.10)$$

Next result follows from (4.2) and Lemmata 4.3 and 4.4.

**Corollary 4.5.** *The sequence of probability measures  $\mu_n$  converges to  $\mu$ .*

Fix  $\mu \in \mathcal{P}(\mathcal{V}^{(p)})$  which can be represented as  $\mu = \sum_{j \in S_p} \omega_j \pi_j^{(p)}$  for some  $\omega \in \mathcal{P}(S_p)$  such that  $\omega_j > 0$  for all  $j \in S_p$ . Recall the definition of  $\Omega_a$ ,  $\omega^{(a)}$  introduced in Lemma 4.3. By definition of  $\mathcal{J}^{(p)}$ , Lemma A.7 and (A.14),

$$\mathcal{J}^{(p)}(\mu) = \sum_{a=1}^{m_p} \Omega_a I_{\mathbb{L}_a^{(p)}}(\omega^{(a)}) + \sum_{a=1}^{\ell_p} \sum_{b \neq a} \sum_{j \in \Omega_a^{(p)}} \sum_{k \in \Omega_b^{(p)}} \omega_j r^{(p)}(j, k). \quad (4.11)$$

Note that we can restrict the first sum in the second term of the right-hand side to the transient equivalent classes.

**Lemma 4.6.** *Fix  $1 \leq p \leq \mathfrak{q}$ , and a measure  $\mu \in \mathcal{P}(\mathcal{V}^{(p)})$  which can be represented as  $\mu = \sum_{j \in S_p} \omega_j \pi_j^{(p)}$  for some  $\omega \in \mathcal{P}(S_p)$  such that  $\omega_j > 0$  for all  $j \in S_p$ . Let  $(\mu_n : n \geq 1)$  be the sequence of probability measures introduced in (4.10). Then,*

$$\limsup_{n \rightarrow \infty} \theta_n^{(p)} \mathcal{J}_n^{(p)}(\mu_n) \leq \mathcal{J}^{(p)}(\mu).$$

*Proof.* By convexity,

$$\mathcal{J}_n^{(p)}(\mu_n) \leq \sum_{a=1}^{m_p} \Omega_a \mathcal{J}_n^{(p)}(\mu_n^{p,a}) + \sum_{j \in S_p^{\text{sgl}}} \omega_j \mathcal{J}_n^{(p)}(\nu_n^{p,j}). \quad (4.12)$$

We investigate the asymptotic behavior of each term separately.

Fix  $j \in S_p^{\text{sgl}}$ . By Lemma A.7 and equation (A.14), since the support of the measure  $\nu_n^{p,j}$  is the set  $\mathcal{V}_j^{(p)}$ , and, by Lemma 3.1, the trace process  $Y_t^{n,p}$  reflected at  $\mathcal{V}_j^{(p)}$  is irreducible,

$$\mathcal{J}_n^{(p)}(\nu_n^{p,j}) = I_{\mathfrak{R}_{\mathcal{V}_j^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n}(\nu_n^{p,j}) + \sum_{x \in \mathcal{V}_j^{(p)}} \sum_{y \in \mathcal{V}^{(p)} \setminus \mathcal{V}_j^{(p)}} \nu_n^{p,j}(x) R_n^{(p)}(x, y).$$

The first term on the right-hand side vanishes because  $\nu_n^{p,j}$  is the stationary state of the process induced by the generator  $\mathfrak{R}_{\mathcal{V}_j^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n$ . By the proof of (4.6), or (4.8), the second term multiplied by  $\theta_n^{(p)}$  converges to  $\sum_{k \in S_p \setminus \{j\}} r^{(p)}(j, k)$  so that

$$\lim_{n \rightarrow \infty} \theta_n^{(p)} \sum_{j \in S_p^{\text{sgl}}} \omega_j \mathcal{J}_n^{(p)}(\nu_n^{p,j}) = \sum_{j \in S_p^{\text{sgl}}} \sum_{k \in S_p \setminus \{j\}} \omega_j r^{(p)}(j, k).$$

The analysis of the asymptotic behavior of the first term on the right-hand side in (4.12) is similar. Fix  $1 \leq a \leq m_p$ . Since the support of the measure  $\mu_n^{p,a}$  is the

set  $\mathcal{W}_a^{(p)}$ , and, by Lemma 4.2, the trace process  $Y_t^{n,p}$  reflected at  $\mathcal{W}_a^{(p)}$  is irreducible, by Lemma (A.7) and equation (A.14),

$$\mathcal{J}_n^{(p)}(\mu_n^{p,a}) = I_{\mathfrak{R}_{\mathcal{W}_a^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n}(\mu_n^{p,a}) + \sum_{x \in \mathcal{W}_a^{(p)}} \sum_{y \in \mathcal{V}^{(p)} \setminus \mathcal{W}_a^{(p)}} \mu_n^{p,a}(x) R_n^{(p)}(x, y). \quad (4.13)$$

Since  $\mu_n^{p,a}$  is the stationary state of the dynamics induced by the generator  $\mathfrak{M}_{H_a}$   $\mathfrak{R}_{\mathcal{W}_a^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n$ , where  $H_a$  is the function introduced in (4.3), by Lemma A.2,

$$I_{\mathfrak{R}_{\mathcal{W}_a^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n}(\mu_n^{p,a}) = - \sum_{x \in \mathcal{W}_a^{(p)}} \sum_{y \in \mathcal{W}_a^{(p)} \setminus \{x\}} \mu_n^{p,a}(x) R_n(x, y) [e^{H_a(y) - H_a(x)} - 1].$$

As  $H_a$  is constant and equal to  $\mathbf{h}_a(j)$  on each set  $\mathcal{V}_j^{(p)}$ ,  $j \in \mathfrak{D}_a^{(p)} = S_p^a$ , the previous expression is equal to

$$- \sum_{j \in S_p^a} \sum_{k \in S_p^a \setminus \{j\}} [e^{\mathbf{h}_a(k) - \mathbf{h}_a(j)} - 1] \sum_{x \in \mathcal{V}_j^{(p)}} \sum_{y \in \mathcal{V}_k^{(p)}} \mu_n^{p,a}(x) R_n(x, y).$$

Hence, by the proof of (4.6), or (4.8),

$$\lim_{n \rightarrow \infty} \theta_n^{(p)} I_{\mathfrak{R}_{\mathcal{W}_a^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n}(\mu_n^{p,a}) = - \sum_{j \in S_p^a} \sum_{k \in S_p^a \setminus \{j\}} [e^{\mathbf{h}_a(k) - \mathbf{h}_a(j)} - 1] \omega_j^{(a)} r^{(p)}(j, k).$$

As  $\mathbf{h}_a: S_p^a \rightarrow \mathbb{R}$  is the function which solves the optimal problem (A.3) for  $I_{\mathbb{L}_a^{(p)}}(\omega^{(a)})$ , the right-hand side is equal to  $I_{\mathbb{L}_a^{(p)}}(\omega^{(a)})$  so that

$$\lim_{n \rightarrow \infty} \theta_n^{(p)} I_{\mathfrak{R}_{\mathcal{W}_a^{(p)}} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n}(\mu_n^{p,a}) = I_{\mathbb{L}_a^{(p)}}(\omega^{(a)}).$$

By similar reasons, the second term on the right-hand in (4.13) multiplied by  $\theta_n^{(p)}$  converges to  $\sum_{j \in S_p^a} \sum_{k \in S_p \setminus S_p^a} \omega_j^{(a)} r^{(p)}(j, k)$  so that

$$\lim_{n \rightarrow \infty} \theta_n^{(p)} \sum_{a=1}^{m_p} \Omega_a \mathcal{J}_n^{(p)}(\mu_n^{p,a}) = \sum_{a=1}^{m_p} \Omega_a I_{\mathbb{L}_a^{(p)}}(\omega^{(a)}) + \sum_{a=1}^{m_p} \sum_{j \in S_p^a} \sum_{k \in S_p \setminus S_p^a} \omega_j r^{(p)}(j, k).$$

By (4.11), the right-hand side is equal to  $\mathcal{J}^{(p)}(\mu)$ , completing the proof of the lemma.  $\square$

*Proof of Proposition 4.1.* Fix  $1 \leq p \leq \mathfrak{q}$  and  $\mu \in \mathcal{P}(\mathcal{V}^{(p)})$ . By Lemmata B.4 and B.5, we may assume that  $\mu(x) > 0$  for all  $x \in \mathcal{V}^{(p)}$ . If  $\mathcal{J}^{(p)}(\mu) = \infty$ , there is nothing to prove. Assume, therefore, that  $\mu = \sum_{j \in S_p} \omega_j \pi_j^{(p)}$  for some  $\omega \in \mathcal{P}(S_p)$ . Since  $\mu(x) > 0$  for all  $x \in \mathcal{V}^{(p)}$ ,  $\omega_j > 0$  for all  $j \in S_p$ . To complete the proof it remains to recall the assertions of Corollary 4.5 and Lemma 4.6.  $\square$

## 5. PROOFS OF PROPOSITION 2.1 AND THEOREM 2.3

We first present some properties of the functionals  $\mathcal{J}^{(p)}$ .

**Lemma 5.1.** *Fix  $1 \leq p < \mathfrak{q}$ . Then,*

$$\mathcal{J}^{(p+1)}(\mu) < \infty \quad \text{if and only if} \quad \mathcal{J}^{(p)}(\mu) = 0.$$

*Proof.* Suppose that  $\mathcal{J}^{(p)}(\mu) = 0$ . Then, by (2.17),  $\mu = \sum_{j \in S_p} \omega_j \pi_j^{(p)}$  for some  $\omega \in \mathcal{P}(S_p)$  and  $\mathbb{I}^{(p)}(\omega) = 0$ . By the definition (2.16) of  $\mathbb{I}^{(p)}$  and Lemma A.8,  $\omega$  is a stationary state of the Markov chain  $\mathbb{X}_t^{(p)}$ , that is,  $\omega$  is a convex combination of the measures  $M_m^{(p)}$ ,  $m \in S_{p+1}$ :

$$\omega(j) = \sum_{m \in S_{p+1}} \vartheta(m) M_m^{(p)}(j), \quad j \in S_p,$$

for some  $\vartheta \in \mathcal{P}(S_{p+1})$ . Inserting this expression in the formula for  $\mu$  and changing the order of summation yields that

$$\mu = \sum_{m \in S_{p+1}} \vartheta(m) \sum_{j \in S_p} M_m^{(p)}(j) \pi_j^{(p)} = \sum_{m \in S_{p+1}} \vartheta(m) \pi_m^{(p+1)},$$

where we used identity (2.12) in the last step. This proves the first assertion of the lemma because  $\mathbb{I}^{(p+1)}(\vartheta) < \infty$  for all  $\vartheta \in \mathcal{P}(S_{p+1})$ . We turn to the converse.

Suppose that  $\mathcal{J}^{(p+1)}(\mu) < \infty$ . In this case, by (2.17),  $\mu = \sum_{m \in S_{p+1}} \vartheta(m) \pi_m^{(p+1)}$  for some  $\vartheta \in \mathcal{P}(S_{p+1})$ . By (2.12), this identity can be rewritten as

$$\mu(\cdot) = \sum_{j \in S_p} \left( \sum_{m \in S_{p+1}} \vartheta(m) M_m^{(p)}(j) \right) \pi_j^{(p)}(\cdot).$$

Therefore, by definition of  $\mathcal{J}^{(p)}$ ,  $\mathcal{J}^{(p)}(\mu) = \mathbb{I}^{(p)}(\omega)$ , where  $\omega(j) = \sum_{m \in S_{p+1}} \vartheta(m) M_m^{(p)}(j)$ .

As the measures  $M_m^{(p)}$  are stationary for the chain  $\mathbb{X}_t^{(p)}$ , so is  $\omega$ . Thus, by Lemma A.8,  $\mathbb{I}^{(p)}(\omega) = 0$ , as claimed.  $\square$

By Lemma A.8,  $\mathcal{J}^{(0)}(\mu) = 0$  if and only if there exists a probability measure  $\omega$  on  $S_1$  such that

$$\mu = \sum_{j \in S_1} \omega_j \pi_j^{(1)}.$$

By (2.17) and since  $\mathbb{I}^{(1)}(\omega) < \infty$  for all  $\omega \in \mathcal{P}(S_1)$ ,  $\mu$  has this form if and only if  $\mathcal{J}^{(1)}(\mu) < \infty$ . Hence, the previous lemma holds for  $p = 0$  as well:

$$\mathcal{J}^{(1)}(\mu) < \infty \quad \text{if and only if} \quad \mathcal{J}^{(0)}(\mu) = 0. \quad (5.1)$$

We turn to the proof of Proposition 2.1. In sake of completeness, we reproduce the proof that  $\mathcal{J}^{(0)}$  is a  $\Gamma$ -lim inf of the sequence  $\mathcal{J}_n$  presented in [5] and which applies to non-reversible dynamics. Next result is the second half of [5, Proposition 8.3].

**Lemma 5.2.** *The functional  $\mathcal{J}^{(0)}: \mathcal{P}(V) \rightarrow \mathbb{R}_+$  is a  $\Gamma$ -lim inf of the sequence  $\mathcal{J}_n$ .*

*Proof.* Fix  $\mu \in \mathcal{P}(V)$  and a sequence of probability measures  $\mu_n$  in  $\mathcal{P}(V)$  converging to  $\mu$ . By definition of  $\mathcal{J}_n$ ,

$$\mathcal{J}_n(\mu_n) \geq - \int_V \frac{\mathcal{L}_n u}{u} d\mu_n = - \sum_{x \in V} \frac{\mu_n(x)}{u(x)} \sum_{y \in V} R_n(x, y) [u(y) - u(x)]$$

for all  $u: V \rightarrow (0, \infty)$ . As  $\mu_n \rightarrow \mu$  and  $R_n \rightarrow \mathbb{R}_0$ , this expression converges to

$$- \sum_{y \neq x \in V} \frac{\mu(x)}{u(x)} \mathbb{R}_0(x, y) [u(y) - u(x)].$$

Therefore,

$$\liminf_{n \rightarrow \infty} \mathcal{J}_n(\mu_n) \geq \sup_{u > 0} - \sum_{y \neq x \in V} \frac{\mu(x)}{u(x)} \mathbb{R}_0(x, y) [u(y) - u(x)] = \mathcal{J}^{(0)}(\mu),$$

which completes the proof of the lemma.  $\square$

We turn to the  $\Gamma$  – lim sup. Fix a measure  $\mu \in \mathcal{P}(V)$  such that  $\mu(x) > 0$  for all  $x \in V$ . Denote by  $\mathfrak{D}_1^{(0)}, \dots, \mathfrak{D}_{m_0}^{(0)}$  the equivalent classes of the chain  $\mathbb{X}_t$  which are not singletons.

By definition, the Markov chain  $\mathbb{X}_t$  reflected at  $\mathfrak{D}_a^{(0)}$ ,  $1 \leq a \leq m_0$ , is irreducible. Denote by  $\mu_{\mathfrak{D}_a^{(0)}}$  the measure  $\mu$  conditioned to  $\mathfrak{D}_a^{(0)}$  defined by equation (A.13). Let  $H_a : \mathfrak{D}_a^{(0)} \rightarrow \mathbb{R}$  be the function given by Lemma A.3 which turns  $\mu_{\mathfrak{D}_a^{(0)}}$  a stationary state for the Markov chain induced by  $\mathfrak{M}_{H_a} \mathfrak{R}_{\mathfrak{D}_a^{(0)}} \mathbb{L}^{(0)}$ , the generator  $\mathfrak{R}_{\mathfrak{D}_a^{(0)}} \mathbb{L}^{(0)}$  tilted by  $H_a$ .

By Corollary A.10,

$$\mathcal{J}^{(0)}(\mu) = - \sum_{a=1}^{m_0} \sum_{x \in \mathfrak{D}_a} \sum_{y \in \mathfrak{D}_a \setminus \{x\}} \mu(x) \mathbb{R}_{H_a}(x, y) + \sum_{x \in V} \sum_{y \in V \setminus \{x\}} \mu(x) \mathbb{R}_0(x, y), \quad (5.2)$$

where  $\mathbb{R}_{H_a}(x, y) = \mathbb{R}_0(x, y) e^{H_a(y) - H_a(x)}$

**Lemma 5.3.** *For all  $\mu \in \mathcal{P}(V)$ ,*

$$\limsup_{n \rightarrow \infty} \mathcal{J}_n(\mu) \leq \mathcal{J}^{(0)}(\mu).$$

*Proof.* Denote by  $X_t^{\mu, n}$  the Markov chain  $X_t^n$  reflected at  $V_\mu$ , and by  $\mathfrak{D}_1^n, \dots, \mathfrak{D}_{m_n}^n$  the equivalent classes of the chain  $X_t^{\mu, n}$  which are not singletons. By assumption (2.5), the sets  $\mathfrak{D}_a$  do not depend on  $n$  and we may remove the index  $n$  from the notation. Moreover, since  $R_n(x, y)$  converges to  $\mathbb{R}_0(x, y)$ , for all  $1 \leq a \leq m_0$ , there exists  $1 \leq b \leq m$  such that  $\mathfrak{D}_a^{(0)} \subset \mathfrak{D}_b$ .

By Lemma A.7,  $\mathcal{J}_n(\mu) = \mathcal{K}_n(\mu)$ , where  $\mathcal{K}_n(\mu)$  is given by (A.15) with the rates  $R$  replaced by  $R_n$ . The functional  $\mathcal{K}_n$  is composed of two terms. The second, as  $n \rightarrow \infty$ , converges to

$$\sum_{x \in V} \sum_{y \in V \setminus \{x\}} \mu(x) \mathbb{R}_0(x, y).$$

We turn to the first term of  $\mathcal{K}_n$ , given by

$$- \sum_{b=1}^m \sum_{x \in \mathfrak{D}_b} \sum_{y \in \mathfrak{D}_b \setminus \{x\}} \mu(x) R_n(x, y) e^{H_n^{(b)}(y) - H_n^{(b)}(x)}, \quad (5.3)$$

where  $H_n^{(b)} : \mathfrak{D}_b \rightarrow \mathbb{R}$  is the function (unique up to an additive constant) which turns  $\mu$  a stationary state for the chain induced by  $\mathfrak{M}_{H_n^{(b)}} \mathfrak{R}_{\mathfrak{D}_b} \mathcal{L}_n$  (the generator  $\mathfrak{R}_{\mathfrak{D}_b} \mathcal{L}_n$  tilted by  $H_n^{(b)}$ ).

Since for all  $1 \leq a \leq m_0$ , there exists  $1 \leq b \leq m$  such that  $\mathfrak{D}_a^{(0)} \subset \mathfrak{D}_b$ , the sum appearing in the previous displayed equation is bounded above by

$$- \sum_{b=1}^m \sum_a \sum_{x \in \mathfrak{D}_a^{(0)}} \sum_{y \in \mathfrak{D}_a^{(0)} \setminus \{x\}} \mu(x) R_n(x, y) e^{H_n^{(b)}(y) - H_n^{(b)}(x)}, \quad (5.4)$$

where the second sum is performed over all  $1 \leq a \leq m_0$  such that  $\mathcal{D}_a^{(0)} \subset \mathcal{D}_b$ .

Fix  $b$  and  $a$  satisfying  $\mathcal{D}_a^{(0)} \subset \mathcal{D}_b$ . By (A.8), there exists a finite constant  $C_n^{(b)}$  such that  $|H_n^{(b)}(y) - H_n^{(b)}(x)| \leq C_n^{(b)}$  for all  $y \neq x \in \mathcal{D}_b$ . Since  $R_n(x, y) \rightarrow \mathbb{R}_0(x, y)$ , and  $\mu(x) > 0$  for all  $x \in \mathcal{D}_a^{(0)}$ , as  $\mathbb{X}_t^\mu$  is irreducible in  $\mathcal{D}_a^{(0)}$ , by Remark A.4, there exists a finite constant  $C_a$ , independent of  $n$ , such that  $|H_n^{(b)}(y) - H_n^{(b)}(x)| \leq C_a$  for all  $y \neq x \in \mathcal{D}_a^{(0)}$ . Therefore, there exists a function  $G_a: \mathcal{D}_a^{(0)} \rightarrow \mathbb{R}$  and a subsequence  $n'$  such that  $H_{n'}(y) - H_{n'}(x) \rightarrow G(y) - G(x)$  for all  $x, y \in \mathcal{D}_a^{(0)}$ .

In conclusion, through a subsequence, (5.4) converges to

$$\begin{aligned} & - \sum_{a=1}^{m_0} \sum_{x \in \mathcal{D}_a^{(0)}} \sum_{y \in \mathcal{D}_a^{(0)} \setminus \{x\}} \mu(x) \mathbb{R}_0(x, y) e^{G_a(y) - G_a(x)} \\ & \leq - \sum_{a=1}^{m_0} \sum_{x \in \mathcal{D}_a^{(0)}} \sum_{y \in \mathcal{D}_a^{(0)} \setminus \{x\}} \mu(x) \mathbb{R}_0(x, y) e^{H_a(y) - H_a(x)}, \end{aligned}$$

where  $H_a$  is the function which appears in (5.2). The inequality holds because for each  $1 \leq a \leq m_0$ ,  $H_a$  is the function which optimises the sum. To complete the proof of the lemma, it remains to collect the previous estimates.  $\square$

Next result is a consequence of the two previous lemmata.

**Corollary 5.4.** *The functional  $\mathcal{J}_n$   $\Gamma$ -converge to  $\mathcal{J}^{(0)}$ .*

**Proof of Theorem 2.3.** The proof is by induction in  $p$ . The case  $p = 0$  is covered by Corollary 5.4. Fix  $1 \leq p \leq \mathfrak{q}$  and assume that the result holds for  $0 \leq p' < p$ . In sake of completeness we reproduce below the proof of the  $\Gamma$ -lim inf taken from [5, Proposition 8.4].

$\Gamma$ -lim inf: For all  $\mu \in \mathcal{P}(V)$  and all sequence of probability measures  $\mu_n \in \mathcal{P}(V)$  such that  $\mu_n \rightarrow \mu$ ,

$$\liminf_{n \rightarrow \infty} \theta_n^{(p)} \mathcal{J}_n^{(p)}(\mu_n) \geq \mathcal{J}^{(p)}(\mu). \quad (5.5)$$

Fix a probability measure  $\mu$  on  $V$  and a sequence  $\mu_n$  converging to  $\mu$ . Suppose that  $\mathcal{J}^{(p-1)}(\mu) > 0$ . In this case, since  $\theta_n^{(p-1)} \mathcal{J}_n$   $\Gamma$ -converges to  $\mathcal{J}^{(p-1)}$  and  $\theta_n^{(p)} / \theta_n^{(p-1)} \rightarrow \infty$ ,

$$\liminf_{n \rightarrow \infty} \theta_n^{(p)} \mathcal{J}_n(\mu_n) = \liminf_{n \rightarrow \infty} \frac{\theta_n^{(p)}}{\theta_n^{(p-1)}} \theta_n^{(p-1)} \mathcal{J}_n(\mu_n) \geq \mathcal{J}^{(p-1)}(\mu) \lim_{n \rightarrow \infty} \frac{\theta_n^{(p)}}{\theta_n^{(p-1)}} = \infty.$$

On the other hand, by Lemma 5.1,  $\mathcal{J}^{(p)}(\mu) = \infty$ . This proves the  $\Gamma$ -lim inf convergence for measures  $\mu$  such that  $\mathcal{J}^{(p-1)}(\mu) > 0$ .

Assume that  $\mathcal{J}^{(p-1)}(\mu) = 0$ . By Lemma 5.1 and (5.1), there exists a probability measure  $\omega$  on  $S_p$  such that  $\mu = \sum_{j \in S_p} \omega_j \pi_j^{(p)}$ . By definition of  $\mathcal{J}_n$ ,

$$\mathcal{J}_n(\mu_n) \geq - \int_V \frac{\mathcal{L}_n u}{u} d\mu_n$$

for all  $u: V \rightarrow (0, \infty)$ .

Fix a function  $h: \mathcal{V}^{(p)} \rightarrow (0, \infty)$  which is constant on each  $\mathcal{V}_j^{(p)}$ ,  $j \in S_p$ :  $h = \sum_{j \in S_p} \mathbf{h}(j) \chi_{\mathcal{V}_j^{(p)}}$ . Let  $u_n: V \rightarrow \mathbb{R}$  be the solution of the Poisson equation (C.5) with  $\mathcal{L} = \mathcal{L}_n$ ,  $\mathcal{A} = \mathcal{V}^{(p)}$  and  $u = h$ . By the representation (C.7), it is clear that  $u_n(x) \in (0, \infty)$  for all  $x \in V$ .



Since  $u_n$  is harmonic on  $V \setminus \mathcal{V}^{(p)}$  and  $u_n = h$  on  $\mathcal{V}^{(p)}$ , by (C.8), the right-hand side of the previous displayed equation with  $u = u_n$  is equal to

$$- \int_{\mathcal{V}^{(p)}} \frac{\mathcal{L}_n u_n}{u_n} d\mu_n = - \int_{\mathcal{V}^{(p)}} \frac{\mathcal{L}_n u_n}{h} d\mu_n = - \int_{\mathcal{V}^{(p)}} \frac{(\mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n) h}{h} d\mu_n .$$

Since  $h$  is constant on each set  $\mathcal{V}_j^{(p)}$  (and equal to  $\mathbf{h}(j)$ ), the last integral is equal to

$$- \sum_{j,k \in S_p} \frac{[\mathbf{h}(k) - \mathbf{h}(j)]}{\mathbf{h}(j)} \sum_{x \in \mathcal{V}_k^{(p)}} \pi_n(x) \frac{\mu_n(x)}{\pi_n(x)} R_n^{(p)}(x, \mathcal{V}_k^{(p)}) ,$$

where  $R_n^{(p)}(x, \mathcal{V}_k^{(p)}) = \sum_{y \in \mathcal{V}_k^{(p)}} R_n^{(p)}(x, y)$ . By (2.13),  $\pi_n(x)/\pi_n(\mathcal{V}_j^{(p)}) \rightarrow \pi_j^{(p)}(x)$  for all  $x \in \mathcal{V}_j^{(p)}$ . Thus, since  $\mu_n \rightarrow \mu = \sum_{j \in S_p} \omega_j \pi_j^{(p)}$ ,

$$\lim_{n \rightarrow \infty} \pi_n(\mathcal{V}_j^{(p)}) \frac{\mu_n(x)}{\pi_n(x)} = \omega_j \quad \text{for all } x \in \mathcal{V}_j^{(p)} .$$

Therefore, by (3.2), (3.3), as  $n \rightarrow \infty$ , the penultimate expression multiplied by  $\theta_n^{(p)}$  converges to

$$- \sum_{j \in S_p} \omega_j \frac{1}{\mathbf{h}(j)} \sum_{k \in S_p} r^{(p)}(j, k) [\mathbf{h}(k) - \mathbf{h}(j)] = - \sum_{j \in S_p} \omega_j \frac{\mathbb{L}^{(p)} \mathbf{h}}{\mathbf{h}} .$$

Summarising, we proved that

$$\liminf_{n \rightarrow \infty} \theta_n^{(p)} \mathcal{J}_n(\mu_n) \geq \sup_{\mathbf{h}} - \sum_{j \in S_p} \omega_j \frac{\mathbb{L}^{(p)} \mathbf{h}}{\mathbf{h}} ,$$

where the supremum is carried over all functions  $\mathbf{h} : S_p \rightarrow (0, \infty)$ . By (2.16), (2.17), the right-hand side is precisely  $\mathcal{J}^{(p)}(\mu)$ , which completes the proof of the  $\Gamma$ -lim inf.  $\Gamma$ -lim sup. Fix  $\mu \in \mathcal{P}(V)$ . If  $\mathcal{J}^{(p)}(\mu) = \infty$ , there is nothing to prove. Assume, therefore, that  $\mu = \sum_{j \in S_p} \omega_j \pi_j^{(p)}$  for some  $\omega \in \mathcal{P}(S_p)$ .

By Lemmata B.4 and B.3, it is enough to prove the theorem for measures  $\mu = \sum_{j \in S_p} \omega_j \pi_j^{(p)}$  for some  $\omega \in \mathcal{P}(S_p)$  such that  $\omega_j > 0$  for all  $j \in S_p$ . Fix such a measure  $\mu$ . Let  $\mu_n \in \mathcal{P}(\mathcal{V}^{(p)})$  be the measure given by (4.10). By Corollary 4.5 and Lemma 4.6,  $\mu_n \rightarrow \mu$  and

$$\limsup_{n \rightarrow \infty} \theta_n^{(p)} \mathcal{J}_n^{(p)}(\mu_n) \leq \mathcal{J}^{(p)}(\mu) . \quad (5.6)$$

Since the trace process  $Y_t^{n,p}$  is irreducible and  $\mu_n(x) > 0$  for all  $x \in \mathcal{V}^{(p)}$ , by Lemma A.3, there exists  $u_n : \mathcal{V}^{(p)} \rightarrow (0, \infty)$  such that

$$\mathcal{J}_n^{(p)}(\mu_n) = - \int_{\mathcal{V}^{(p)}} \frac{1}{u_n} [(\mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n) u_n] d\mu_n .$$

Denote by  $v_n$  the harmonic extension of  $u_n$  to  $V$  given by (C.5) with  $\mathcal{A}, \mathcal{L}$  replaced by  $\mathcal{V}^{(p)}, \mathcal{L}_n$ , respectively. Let  $\nu_n$  be the stationary state of the tilted generator  $\mathfrak{M}_{v_n} \mathcal{L}_n$ . By Proposition C.1,

$$\mathcal{J}_n(\nu_n) \leq \mathcal{J}_n^{(p)}(\mu_n) . \quad (5.7)$$

In view of (5.6), (5.7), it remains to show that  $\nu_n \rightarrow \mu$ .

As  $\nu_n$  is the stationary state of the Markov chain  $X_t^{(n)}$  tilted by  $v_n$ , by [2, Proposition 6.3],  $\nu_n$  conditioned to  $\mathcal{V}^{(p)}$  is the stationary state of the Markov chain

induced by the generator  $\mathfrak{T}_{\mathcal{V}^{(p)}} \mathfrak{M}_{\nu_n} \mathcal{L}_n$ . By Lemma C.4 this generator coincides with  $\mathfrak{M}_{\nu_n} \mathfrak{T}_{\mathcal{V}^{(p)}} \mathcal{L}_n$ . By definition,  $\mu_n$  is the stationary state of this later Markov chain. Hence,  $\mu_n(\cdot) = \nu_n(\cdot | \mathcal{V}^{(p)})$ .

Since  $\mu_n \rightarrow \mu$  and  $\mu_n(\cdot) = \nu_n(\cdot | \mathcal{V}^{(p)})$ , it is enough to show that  $\nu_n(\mathcal{V}^{(p)}) \rightarrow 1$ . Assume, by contradiction, that  $\limsup_n \nu_n(z) > 0$  for some  $z \in V \setminus \mathcal{V}^{(p)}$ . Since  $\mathcal{P}(V)$  is compact for the weak topology, consider a subsequence, still denoted by  $\nu_n$ , such that  $\nu_n \rightarrow \nu \in \mathcal{P}(V)$ ,  $\nu(z) > 0$ . By the  $\Gamma$ -lim inf,

$$\liminf_{n \rightarrow \infty} \theta_n^{(p)} \mathcal{J}_n(\nu_n) \geq \mathcal{J}^{(p)}(\nu).$$

Since  $\nu(z) > 0$ ,  $z \in V \setminus \mathcal{V}^{(p)}$ ,  $\mathcal{J}^{(p)}(\nu) = +\infty$ . However, by (5.6), (5.7),

$$\limsup_{n \rightarrow \infty} \theta_n^{(p)} \mathcal{J}_n(\nu_n) \leq \mathcal{J}^{(p)}(\mu) = \mathbb{I}^{(p)}(\omega) < \infty.$$

Hence,  $\nu_n(\mathcal{V}^{(p)}) \rightarrow 1$  and  $\nu_n \rightarrow \mu$ , which completes the proof of the theorem.  $\square$

#### APPENDIX A. THE RATE FUNCTION

Fix a finite set  $V$ . Consider a  $V$ -valued continuous-time Markov chain  $(X_t : t \geq 0)$ , and denote by  $\mathcal{L}$  its generator. The jump rates are represented by  $R(\cdot, \cdot)$ , so that

$$(\mathcal{L}f)(x) = \sum_{y \in V} R(x, y) \{ f(y) - f(x) \} \quad (\text{A.1})$$

for all functions  $f: V \rightarrow \mathbb{R}$ . Denote by  $\lambda(x)$ ,  $x \in V$ , the holding rates and by  $p(x, y)$ ,  $x, y \in V$ , the jump probabilities, so that  $R(x, y) = \lambda(x) p(x, y)$ . Mind that we do not suppose the process to be irreducible. We assume, however, that  $\lambda(x) > 0$  for all  $x \in V$ . The case where some holding rates might vanish is considered at the end of this section.

Denote by  $\mathcal{P}(V)$  the space of probability measures on  $V$  endowed with the weak topology. For a function  $H: V \rightarrow \mathbb{R}$ , define  $J_H: \mathcal{P}(V) \rightarrow \mathbb{R}$  by

$$J_H(\mu) := - \int_V e^{-H} \mathcal{L} e^H d\mu = - \sum_{x, y \in V} \mu(x) R(x, y) [e^{H(y)-H(x)} - 1], \quad (\text{A.2})$$

and let

$$I(\mu) := \sup_H J_H(\mu), \quad (\text{A.3})$$

where the supremum is carried over all functions  $H: V \rightarrow \mathbb{R}$ .

To stress the dependence of the functionals  $J_H$  and  $I$  on the generator  $\mathcal{L}$ , we sometimes denote them by  $J_{\mathcal{L}, H}$  and  $I_{\mathcal{L}}$ , respectively. Next result collects simple properties of the functionals  $J_H$  and  $I$ .

**Lemma A.1.** *For each  $\mu \in \mathcal{P}(V)$ , the functional  $H \mapsto J_H(\mu)$  is concave, and  $J_{H+c}(\mu) = J_H(\mu)$  for all constants  $c$ . The functional  $I$  is convex, lower-semicontinuous, non-negative, and bounded by  $\sum_{x \in V} \mu(x) \lambda(x)$ .*

Recall that a measure  $\mu \in \mathcal{P}(V)$  is a stationary state for the Markov chain induced by the generator  $\mathcal{L}$  if  $\int_V (\mathcal{L}f) d\mu = 0$  for all function  $f: V \rightarrow \mathbb{R}$ . As we do not assume the chain to be irreducible, the stationary state may not be unique.

The Euler-Lagrange equation for  $I$  reads as

$$\int_V (\mathfrak{M}_H \mathcal{L}) G d\mu = 0 \quad \text{for all } G: V \rightarrow \mathbb{R}. \quad (\text{A.4})$$

In this formula, for a function  $H: V \rightarrow \mathbb{R}$ ,  $\mathfrak{M}_H \mathcal{L}$  represents the tilted generator given by

$$[(\mathfrak{M}_H \mathcal{L}) f](x) = \sum_{y \in V} e^{-H(x)} R(x, y) e^{H(y)} [f(y) - f(x)] \quad (\text{A.5})$$

for  $f: V \rightarrow \mathbb{R}$ . Let  $R_H(x, y) := e^{-H(x)} R(x, y) e^{H(y)}$ . Next result clarify the meaning of the Euler-Lagrange equation (A.4).

**Lemma A.2.** *A probability measure  $\mu$  in  $V$  is a stationary state for the Markov chain induced by the generator  $\mathfrak{M}_H \mathcal{L}$  if, and only if,*

$$I(\mu) = J_H(\mu). \quad (\text{A.6})$$

*Proof.* Suppose that  $\mu$  is a stationary state for the Markov chain induced by the generator  $\mathfrak{M}_H \mathcal{L}$ . Then, for all functions  $G: V \rightarrow \mathbb{R}$ ,

$$\begin{aligned} J_G(\mu) - J_H(\mu) &= - \sum_{x, y \in V} \mu(x) R(x, y) [e^{G(y) - G(x)} - e^{H(y) - H(x)}] \\ &= - \sum_{x, y \in V} \mu(x) R_H(x, y) [e^{F(y) - F(x)} - 1], \end{aligned}$$

where  $F = G - H$ . Since  $\mu$  is a stationary state for the Markov chain induced by the generator  $\mathfrak{M}_H \mathcal{L}$ , the previous expression is equal to

$$- \sum_{x, y \in V} \mu(x) R_H(x, y) \{ e^{F(y) - F(x)} - 1 - [F(y) - F(x)] \}.$$

This expression is negative because  $a \mapsto e^a - 1 - a$  is positive. This proves that  $I(\mu) = \sup_G J_G(\mu) \leq J_H(\mu)$ , as claimed.

Conversely, suppose that (A.6) holds. Fix  $G: V \rightarrow \mathbb{R}$ . Then, the function  $a \mapsto J_{H+aG}(\mu)$  assumes a maximum at  $a = 0$ . Its derivative at  $a = 0$  is given by  $\int_V (\mathfrak{M}_H \mathcal{L}) G d\mu$ . Hence, (A.4) holds for all  $G$  yielding that  $\mu$  is stationary for the Markov chain induced by the generator  $\mathfrak{M}_H \mathcal{L}$ .  $\square$

**Lemma A.3.** *Assume that the Markov chain induced by the generator  $\mathcal{L}$  is irreducible and that  $\mu(x) > 0$  for all  $x \in V$ . Then, there exists a function  $H: V \rightarrow \mathbb{R}$ , unique up to an additive constant, such that*

$$I(\mu) = J_H(\mu).$$

Moreover,

$$I(\mu) = \sum_{x, y \in V} \mu(x) R(x, y) \left\{ [H(y) - H(x)] e^{H(y) - H(x)} - e^{H(y) - H(x)} + 1 \right\}, \quad (\text{A.7})$$

and

$$\max_{x, y \in V} |H(y) - H(x)| \leq |V| \ln \frac{1 + \sum_{x, y \in V} \mu(x) R(x, y)}{\min_{z, w} \mu(z) R(z, y)}, \quad (\text{A.8})$$

where the minimum is performed over all edges  $(z, w)$  such that  $R(z, w) > 0$ .

*Proof.* Since  $I(\mu)$  is bounded, there exists a sequence  $(H_n: n \geq 1)$  of functions  $H_n: V \rightarrow \mathbb{R}$  such that

$$I(\mu) = \lim_{n \rightarrow \infty} J_{H_n}(\mu).$$

Fix  $x_0 \in V$ . Since  $J_{H+c}(\mu) = J_H(\mu)$ , redefine the sequence  $H_n$  so that  $H_n(x_0) = 0$  for all  $n \geq 1$ .

*Claim 1:* The sequence  $H_n$  is uniformly bounded.

By definition of the sequence  $H_n$  and since  $I$  is positive, there exists  $n_0 \geq 1$  such that

$$\sum_{x,y \in V} \mu(x) R(x,y) [e^{H_n(y)-H_n(x)} - 1] \leq 1$$

for all  $n \geq n_0$ . Hence,

$$\sum_{x,y \in V} \mu(x) R(x,y) e^{H_n(y)-H_n(x)} \leq C_0 := 1 + \sum_{x,y \in V} \mu(x) R(x,y).$$

Thus,

$$H_n(y) - H_n(x) \leq C_1 := \ln \frac{C_0}{\min_{x,y} \mu(x) R(x,y)} \quad (\text{A.9})$$

for all  $n \geq n_0$  and all edges  $(x,y)$  such that  $R(x,y) > 0$ . In this equation the minimum is performed over all edges  $(x,y)$  such that  $\mu(x) R(x,y) > 0$ .

Fix  $x \in V \setminus \{x_0\}$ . Since the process is irreducible, there exists a self-avoiding path  $x_0, x_1, \dots, x_k = x$  such that  $R(x_i, x_{i+1}) > 0$  for all  $0 \leq i < k$ . Hence, by (A.9),  $H_n(x) = H_n(x_k) - H_n(x_0) \leq C_1 k \leq C_2 := C_1 |V|$ .

Conversely, there exists a self-avoiding path  $x = y_0, y_1, \dots, y_j = x_0$  such that  $R(y_i, y_{i+1}) > 0$  for all  $0 \leq i < j$ . Hence, by (A.9),  $-H_n(x) = H_n(x_0) - H_n(x) = H_n(y_j) - H_n(y_0) \leq C_1 j \leq C_1 |V| = C_2$ , which proves Claim 1.

As the sequence  $H_n$  is uniformly bounded, we may extract a subsequence, still denoted by  $H_n$ , which converges pointwisely to a function  $H$ . By definition of the sequence  $H_n$  and by continuity,

$$I(\mu) = \lim_{n \rightarrow \infty} J_{H_n}(\mu) = J_H(\mu),$$

as asserted. Moreover,  $\max_{x,y \in V} |H(y) - H(x)| \leq C_2$ , proving (A.8).

We turn to the proof of uniqueness. Assume that there are two functions, denoted by  $H$  and  $G$ , which minimize. Let  $F_\theta = \theta H + (1 - \theta)G$ ,  $0 \leq \theta \leq 1$ . By concavity of the functional  $J$ , for all  $0 \leq \theta \leq 1$ ,

$$I(\mu) \geq J_{F_\theta}(\mu) \geq \theta J_H(\mu) + (1 - \theta) J_G(\mu) = I(\mu).$$

Hence,  $\theta \mapsto J_{F_\theta}(\mu)$  is constant. Taking the second derivative yields that

$$\sum_{x,y \in V} \mu(x) R(x,y) \{ [G(y) - G(x)] - [H(y) - H(x)] \}^2 e^{F_\theta(y)-F_\theta(x)} = 0$$

for all  $0 < \theta < 1$ . Hence,  $G(y) - G(x) = H(y) - H(x)$  if  $\mu(x) R(x,y) > 0$ . As the process is irreducible and the measure positive,  $G = H + c$  for some constant  $c \in \mathbb{R}$ .

To show the validity of (A.7), note that

$$I(\mu) = J_H(\mu) = \sum_{x,y \in V} \mu(x) R(x,y) \{ 1 - e^{H(y)-H(x)} \}.$$

Since, by Lemma A.2,  $\mu$  is a stationary state for the Markov chain induced by the generator  $\mathfrak{M}_H \mathcal{L}$ ,

$$0 = \sum_{x,y \in V} \mu(x) R(x,y) e^{H(y)-H(x)} [H(y) - H(x)].$$

Adding the two previous identities yields (A.7).  $\square$

**Remark A.4.** *It follows from the previous proof that the minimum in the denominator in equation (A.8) can be restricted to a subset of edges  $E_0$  which keeps the chain irreducible.*

Next result follows from the two previous lemmata.

**Corollary A.5.** *Assume that the Markov chain induced by the generator  $\mathcal{L}$  is irreducible and that  $\mu(x) > 0$  for all  $x \in V$ . Then, there exists a function  $H: V \rightarrow \mathbb{R}$  such that  $\mu$  is stationary for  $\mathfrak{M}_H \mathcal{L}$ .*

**Corollary A.6.** *Assume that the Markov chain induced by the generator  $\mathcal{L}$  is irreducible and that  $\mu(x) > 0$  for all  $x \in V$ . Then,  $I(\mu) = 0$  if and only if  $\mu$  is the stationary state.*

*Proof.* Assume that  $I(\mu) = 0$ . Then, since  $a e^a - e^a + 1 \geq 0$ , each term in the sum (A.7) vanishes, and  $H(y) = H(x)$  if  $\mu(x) R(x, y) > 0$ . As the Markov chain is irreducible,  $H$  is constant. To complete the argument, it remains to recall that, by Lemma A.2,  $\mu$  is a stationary state for the chain induced by the generator  $\mathfrak{M}_H \mathcal{L} = \mathcal{L}$ .

Conversely, suppose that  $\mu$  is the stationary state. Then,  $\mu$  is a stationary state for the chain induced by the generator  $\mathfrak{M}_H \mathcal{L}$  for  $H = 0$ . By Lemma A.2,  $I(\mu) = J_0(\mu) = 0$ , as claimed.  $\square$

**Reducible Markov chains.** In this subsection, we derive a formula for  $I(\mu)$  in the case where the process  $X_t$  is reducible or the support of  $\mu$  a proper subset of  $V$ .

Denote by  $\mathfrak{R}_{\mathcal{A}} \mathcal{L}$ ,  $\mathcal{A}$  a proper subset of  $V$  which is not a singleton, the generator of the Markov chain  $X_t$  reflected at  $\mathcal{A}$ . This is the  $\mathcal{A}$ -valued Markov chain which jumps from  $x \in \mathcal{A}$  to  $y \in \mathcal{A}$  at rate  $R(x, y)$ . Its generator reads as

$$[(\mathfrak{R}_{\mathcal{A}} \mathcal{L}) f](x) = \sum_{y \in \mathcal{A}} R(x, y) [f(y) - f(x)], \quad x \in \mathcal{A}. \quad (\text{A.10})$$

Clearly, this chain may be reducible even if the original one is irreducible.

Fix a probability measure  $\mu \in \mathcal{P}(V)$ , and denote by  $V_\mu$  its support,  $V_\mu = \{x \in V : \mu(x) > 0\}$ , and by  $X_t^\mu$  the Markov chain reflected at  $V_\mu$ . Mind that we do not assume  $V_\mu$  to be a proper subset of  $V$ .

The formula for  $I(\mu)$  relies on the construction of a directed graph without directed loops. Denote by  $\mathcal{Q}_1, \dots, \mathcal{Q}_\ell$  the equivalence classes of the chain  $X_t^\mu$ . These classes form the set of vertices of the directed graph. Draw a directed arrow from  $\mathcal{Q}_a$  to  $\mathcal{Q}_b$  if there exists  $x \in \mathcal{Q}_a$  and  $y \in \mathcal{Q}_b$  such that  $R(x, y) > 0$ . Denote the set of directed edges by  $\mathbb{A}$  and the graph by  $\mathbb{G} = (\mathbb{Q}, \mathbb{A})$ , where  $\mathbb{Q}$  is the set  $\{\mathcal{Q}_1, \dots, \mathcal{Q}_\ell\}$  of vertices.

A path in the graph  $\mathbb{G}$  is a sequence vertices  $(\mathcal{Q}_{a_j} : 0 \leq j \leq m)$ , such that there is a directed arrow from  $\mathcal{Q}_{a_j}$  to  $\mathcal{Q}_{a_{j+1}}$  for  $0 \leq j < m$ . This directed graph has no directed loops because the existence of a directed loop would contradict the definition of the sets  $\mathcal{Q}_a$  as equivalent classes. (Mind that undirected loops might exist).

Let  $\mathcal{C}_1, \dots, \mathcal{C}_p$  be the closed irreducible classes and  $\mathcal{J}_1, \dots, \mathcal{J}_q$  be the transient ones, so that  $p + q = \ell$ . Since the sets  $\mathcal{C}_j$  are closed irreducible classes, these sets are not the tail of a directed edge in the graph. On the other hand, as the elements of  $\mathcal{J}_i$  are transient for the chain  $X_t^\mu$ , there is a path  $(\mathcal{J}_i = \mathcal{J}_{a_0}, \dots, \mathcal{J}_{a_{m-1}}, \mathcal{C}_j)$  from  $\mathcal{J}_i$  to some irreducible class  $\mathcal{C}_j$ .

Fix a transient class  $\mathcal{T}_i$ . Denote by  $\mathbf{D}(\mathcal{T}_i)$  the length of the longest path from  $\mathcal{T}_i$  to a closed irreducible class. The function  $\mathbf{D}$  is well defined because (a) the set of vertices is finite, (b) there is at least a path, (c) there are no directed loops in the graph.

Fix  $a, b$  such that there is a directed arrow from  $\mathcal{T}_a$  to  $\mathcal{T}_b$ . Then,

$$\mathbf{D}(\mathcal{T}_a) \geq \mathbf{D}(\mathcal{T}_b) + 1. \quad (\text{A.11})$$

Indeed, it is enough to consider the longest path from  $\mathcal{T}_b$  to the irreducible classes.  $\mathcal{T}_a$  does not belong to the path because there are no directed loops. By adding  $\mathcal{T}_a$  at the beginning of the path from  $\mathcal{T}_b$  to the irreducible classes, we obtain a path from  $\mathcal{T}_a$  to the irreducible classes of length  $\mathbf{D}(\mathcal{T}_b) + 1$ , proving (A.11).

Setting  $\mathbf{D}(\mathcal{C}_j) = 0$  for all  $1 \leq j \leq p$ , we may extend (A.11) to the closed irreducible classes. Fix  $a, b$  such that there is a directed arrow from  $\mathcal{T}_a$  to  $\mathcal{C}_b$ . Then,

$$\mathbf{D}(\mathcal{T}_a) \geq \mathbf{D}(\mathcal{C}_b) + 1 \quad (\text{A.12})$$

because  $\mathbf{D}(\mathcal{T}_a) \geq 1$ . Finally, we may lift the function  $\mathbf{D}$  to  $V_\mu$  by setting  $\mathbf{D}(x) = \mathbf{D}(\mathcal{Q}_a)$  for all  $x \in \mathcal{Q}_a$ .

Recall from (A.10) that we represent by  $\mathfrak{R}_{\mathcal{A}}\mathcal{L}$  the generator of the process  $X_t$  reflected at  $\mathcal{A}$ . Assume that the chain induced by the generator  $\mathfrak{R}_{\mathcal{A}}\mathcal{L}$  is irreducible. Let  $I_{\mathfrak{R}_{\mathcal{A}}\mathcal{L}}: \mathcal{P}(\mathcal{A}) \rightarrow \mathbb{R}_+$  the functional given by

$$I_{\mathfrak{R}_{\mathcal{A}}\mathcal{L}}(\mu) := \sup_H - \int_{\mathcal{A}} e^{-H} [(\mathfrak{R}_{\mathcal{A}}\mathcal{L}) e^H] d\mu,$$

where the supremum is carried over all functions  $H: \mathcal{A} \rightarrow \mathbb{R}$ .

Denote by  $\mathcal{D}_a$ ,  $1 \leq a \leq m$  the equivalent classes of the chain  $X_t^\mu$  with at least two elements. Note that  $\mu(\mathcal{D}_a) > 0$  for all  $a$  and that  $m \leq \ell$ . Let  $\mu_{\mathcal{A}}$ ,  $\mathcal{A} \subset V$  such that  $\mu(\mathcal{A}) > 0$ , be the measure  $\mu$  conditioned to  $\mathcal{A}$ :

$$\mu_{\mathcal{A}}(x) := \frac{\mu(x)}{\mu(\mathcal{A})}, \quad x \in \mathcal{A}. \quad (\text{A.13})$$

Let  $K: \mathcal{P}(V) \rightarrow \mathbb{R}_+$  the functional given by

$$\begin{aligned} K(\mu) &= \sum_{a=1}^m \mu(\mathcal{D}_a) I_{\mathfrak{R}_{\mathcal{D}_a}\mathcal{L}}(\mu_{\mathcal{D}_a}) + \sum_{x \in V_\mu} \sum_{y \notin V_\mu} \mu(x) R(x, y) \\ &+ \sum_{a=1}^{\ell} \sum_{b \neq a} \sum_{x \in \mathcal{Q}_a} \sum_{y \in \mathcal{Q}_b} \mu(x) R(x, y). \end{aligned} \quad (\text{A.14})$$

In this formula, since closed irreducible equivalent classes are not the tail of any directed edge, we may restrict the sum over  $a$  to transient equivalent classes. Moreover, if  $R(x, y) > 0$  for some  $x \in \mathcal{Q}_a$ ,  $y \in \mathcal{Q}_b$ , then  $R(z, w) = 0$  for all  $z \in \mathcal{Q}_b$ ,  $w \in \mathcal{Q}_a$ . Hence, in the last sum, each pair  $(a, b)$  is counted only once.

In view of Lemma A.3, we may rewrite (A.14) as

$$K(\mu) = - \sum_{a=1}^m \sum_{x \in \mathcal{D}_a} \sum_{y \in \mathcal{D}_a \setminus \{x\}} \mu(x) R_{H_a}(x, y) + \sum_{x \in V} \sum_{y \in V \setminus \{x\}} \mu(x) R(x, y), \quad (\text{A.15})$$

where  $H_a: \mathcal{D}_a \rightarrow \mathbb{R}$  is the function (unique up to an additive constant) which turns the measure  $\mu$  conditioned to  $\mathcal{D}_a$  stationary for the chain induced by  $\mathfrak{M}_{H_a}\mathfrak{R}_{\mathcal{D}_a}\mathcal{L}$  (the generator  $\mathfrak{R}_{\mathcal{D}_a}\mathcal{L}$  tilted by  $H_a$ ).

**Lemma A.7.** *For all  $\mu \in \mathcal{P}(V)$*

$$I(\mu) = K(\mu) .$$

*Proof.* We first prove that  $I(\mu) \leq K(\mu)$ . Fix  $H: V \rightarrow \mathbb{R}$ . By definition of  $V_\mu$ ,

$$\int_V e^{-H} \mathcal{L} e^H d\mu = \sum_{x \in V_\mu, y \in V} \mu(x) R(x, y) [e^{H(y)-H(x)} - 1] .$$

The right-hand side can be rewritten as

$$\sum_{x \in V_\mu, y \notin V_\mu} c_H(x, y) + \sum_{a=1}^m \sum_{x \in \mathcal{D}_a, y \in \mathcal{D}_a} c_H(x, y) + \sum_{a=1}^\ell \sum_{b \neq a} \sum_{x \in \mathcal{Q}_a} \sum_{y \in \mathcal{Q}_b} c_H(x, y) , \quad (\text{A.16})$$

where  $c_H(x, y) = \mu(x) R(x, y) [\exp\{H(y) - H(x)\} - 1]$ .

As  $c_H(x, y) \geq -\mu(x) R(x, y)$ , the sum of the first and third terms of the previous displayed equation are bounded below by

$$- \sum_{x \in V_\mu, y \notin V_\mu} \mu(x) R(x, y) - \sum_{a=1}^\ell \sum_{b \neq a} \sum_{x \in \mathcal{Q}_a} \sum_{y \in \mathcal{Q}_b} \mu(x) R(x, y) .$$

The second term of that formula is bounded below by

$$\sum_{a=1}^m \inf_G \sum_{x \in \mathcal{D}_a, y \in \mathcal{D}_a} c_G(x, y) = - \sum_{a=1}^m \mu(\mathcal{D}_a) I_{\mathfrak{R}_{\mathcal{D}_a} \mathcal{L}}(\mu_{\mathcal{D}_a}) .$$

Up to this point we proved that

$$\int_V e^{-H} \mathcal{L} e^H d\mu \geq -K(\mu)$$

for all  $H: V \rightarrow \mathbb{R}$ . Multiplying by  $-1$  and optimising over  $H$  yields that  $I(\mu) \leq K(\mu)$ .

We turn to the converse inequality. For each set  $\mathcal{D}_a$  the Markov chain induced by the generator  $\mathfrak{R}_{\mathcal{D}_a} \mathcal{L}$  is irreducible and  $\mu(x) > 0$  for all  $x \in \mathcal{D}_a$ . Hence, by Lemma A.3, there exists a function  $G_a: \mathcal{D}_a \rightarrow \mathbb{R}$  which solves the variational problem

$$I_{\mathfrak{R}_{\mathcal{D}_a} \mathcal{L}}(\mu_{\mathcal{D}_a}) = \sup_H - \int_{\mathcal{D}_a} e^{-H} (\mathfrak{R}_{\mathcal{D}_a} \mathcal{L}) e^H d\mu_{\mathcal{D}_a} = - \int_{\mathcal{D}_a} e^{-G_a} (\mathfrak{R}_{\mathcal{D}_a} \mathcal{L}) e^{G_a} d\mu_{\mathcal{D}_a} , \quad (\text{A.17})$$

where the supremum is carried over all functions  $H: \mathcal{D}_a \rightarrow \mathbb{R}$ .

Recall the definition of the function  $\mathbf{D}: V_\mu \rightarrow \mathbb{R}$  introduced above (A.11). Define the sequence of functions  $H_n: V \rightarrow \mathbb{R}$  by

$$H_n(x) = \begin{cases} -n & x \notin V_\mu , \\ G_a(x) + n \mathbf{D}(x) & x \in \mathcal{D}_a , \\ n \mathbf{D}(x) & x \in V_\mu \setminus \cup_{1 \leq a \leq m} \mathcal{D}_a . \end{cases} \quad (\text{A.18})$$

By definition of the functional  $I$  and since  $V_\mu$  stands for the support of  $\mu$ ,

$$I(\mu) \geq - \liminf_{n \rightarrow \infty} \int_{V_\mu} e^{-H_n} \mathcal{L} e^{H_n} d\mu .$$

For a fixed  $n$  the previous integral is equal to the sum in (A.16) with  $H_n$  replacing  $H$ . By definition of  $H_n$ , as  $n \rightarrow \infty$ , the first term in (A.16) converges to

$$- \sum_{x \in V_\mu, y \notin V_\mu} \mu(x) R(x, y).$$

Since  $\mathbf{D}(x) = \mathbf{D}(y)$  for elements  $x, y$  in the same equivalent class  $\mathcal{D}_a$ , by (A.17), for every  $n \geq 1$ , the second term in (A.16) is equal to

$$\begin{aligned} \sum_{a=1}^m \sum_{x \in \mathcal{D}_a, y \in \mathcal{D}_a} c_{G_a}(x, y) &= \sum_{a=1}^m \mu(\mathcal{D}_a) \int_{\mathcal{D}_a} e^{-G_a} (\mathfrak{R}_{\mathcal{D}_a} \mathcal{L}) e^{G_a} d\mu_{\mathcal{D}_a} \\ &= - \sum_{a=1}^m \mu(\mathcal{D}_a) I_{\mathfrak{R}_{\mathcal{D}_a} \mathcal{L}}(\mu_{\mathcal{D}_a}). \end{aligned}$$

Finally, to estimate the third term in (A.16), fix  $x \in \mathcal{Q}_a, y \in \mathcal{Q}_b$  such that  $R(x, y) > 0$ . By (A.11), (A.12),  $\mathbf{D}(x) \geq \mathbf{D}(y) + 1$ . Thus,  $H_n(y) - H_n(x) \leq -n + C_0$  for some finite constant  $C_0$  independent of  $n$ , and, as  $n \rightarrow \infty$ , the second term in (A.16) converges to

$$- \sum_{a=1}^{\ell} \sum_{b \neq a} \sum_{x \in \mathcal{Q}_a} \sum_{y \in \mathcal{Q}_b} \mu(x) R(x, y)$$

Collecting all previous estimates yields that

$$I(\mu) \geq - \liminf_{n \rightarrow \infty} \int_V e^{-H_n} \mathcal{L} e^{H_n} d\mu = K(\mu),$$

which completes the proof of the lemma.  $\square$

**Lemma A.8.** *A measure  $\mu \in \mathcal{P}(V)$  is a stationary state of the Markov chain  $X_t$  if and only if  $I(\mu) = 0$ .*

*Proof.* Assume that  $I(\mu) = 0$ . Then, by Lemma A.7, all terms on the right-hand side of (A.14) vanish. As the second and third terms vanish the support of  $\mu$  consists of the union of closed irreducible sets of the Markov chain. Since the first term vanishes, by Corollary A.6,  $\mu$  restricted to these irreducible classes is a stationary state. Hence,  $\mu$  is a convex combination of the stationary states, and, hence, a stationary state.

Suppose that  $\mu$  is a stationary state. Then, its support is the union of closed irreducible classes. Therefore, the second and third terms on the right-hand side of (A.14) vanish. Fix a closed irreducible class of the chain contained in the support of  $\mu$ . The restriction of  $\mu$  to this set is strictly positive. Hence, by Corollary A.6, the first term on the right-hand side of (A.14) also vanishes. This completes the proof of the lemma.  $\square$

**Degenerate generators.** In this subsection, we consider generators whose holding rates might vanish. Let  $V_0 = \{x : \lambda(x) > 0\}$  and keep in mind that  $V_0$  may be a proper subset of  $V$ .

Denote by  $\mathcal{L}_0$  the generator  $\mathcal{L}$  restricted to  $V_0$ :

$$(\mathcal{L}_0 f)(x) = \sum_{y \in V_0} R(x, y) \{f(y) - f(x)\}, \quad f: V_0 \rightarrow \mathbb{R}.$$

Fix a measure  $\mu \in \mathcal{P}(V)$ . Clearly, if  $\mu(V_0) = 0$ , then for all  $H: V \rightarrow \mathbb{R}$ ,  $J_H(\mu) = 0$  and  $I(\mu) = 0$ . The next lemma covers the case where  $\mu(V_0) > 0$ . To stress the



dependence of the functional  $J_H$  introduced at (A.2) on the generator  $\mathcal{L}$ , denote it below by  $J_{\mathcal{L},H}$ . For a function  $G: V_0 \rightarrow \mathbb{R}$ , let  $J_{\mathcal{L}_0,G}: \mathcal{P}(V_0) \rightarrow \mathbb{R}$  be the functional given by

$$J_{\mathcal{L}_0,G}(\nu) := - \int_{V_0} e^{-G} \mathcal{L}_0 e^G d\nu .$$

Let  $I_{\mathcal{L}_0}: \mathcal{P}(V_0) \rightarrow \mathbb{R}$  be the functional defined by (A.3) with  $J_H$  replaced by  $J_{\mathcal{L}_0,G}$  and where the supremum is carried over all functions  $G: V_0 \rightarrow \mathbb{R}$ .

**Lemma A.9.** *For all measures  $\mu \in \mathcal{P}(V)$  such that  $\mu(V_0) > 0$ ,*

$$J_{\mathcal{L},H}(\mu) = \mu(V_0) \left\{ J_{\mathcal{L}_0,H_{V_0}}(\mu_{V_0}) - \sum_{x \in V_0} \sum_{y \in V \setminus V_0} \mu_{V_0}(x) R(x,y) [e^{H(y)-H(x)} - 1] \right\} ,$$

where  $H_{V_0}: V_0 \rightarrow \mathbb{R}$  stands for the restriction of  $H$  to  $V_0$ :  $H_{V_0}(x) = H(x)$ ,  $x \in V_0$ . In particular,

$$I_{\mathcal{L}}(\mu) = \mu(V_0) I_{\mathcal{L}_0}(\mu_{V_0}) + \sum_{x \in V_0} \sum_{y \in V \setminus V_0} \mu(x) R(x,y) .$$

*Proof.* The first assertion of the lemma is a simple identity. The proof of the second one is similar to the one of Lemma A.7. We argue as in this lemma to show that  $I_{\mathcal{L}}(\mu)$  is bounded by the right-hand side of the identity. The converse inequality is obtained by observing that to optimize  $J_{\mathcal{L},H}(\mu)$  it is convenient to set  $H(y) = -\infty$  for  $y \in V \setminus V_0$ . More precisely, one proceeds just as in the proof of Lemma A.7 to obtain an optimal sequence  $H_n$  defined in  $V_0$  and then extend it to  $V \setminus V_0$  in such a way that  $H_n(y) - H_n(x) \rightarrow -\infty$  for all  $y \in V \setminus V_0$ ,  $x \in V_0$ .  $\square$

The previous lemma allows us to restrict our attention to non-singular generators. This is the content of the next result.

**Corollary A.10.** *For all measures  $\mu \in \mathcal{P}(V)$  such that  $\mu(V_0) > 0$ ,*

$$I_{\mathcal{L}}(\mu) = - \sum_{a=1}^m \sum_{x \in \mathcal{D}_a} \sum_{y \in \mathcal{D}_a \setminus \{x\}} \mu(x) R_{H_a}(x,y) + \sum_{x \in V} \sum_{y \in V \setminus \{x\}} \mu(x) R(x,y) ,$$

where  $\mathcal{D}_a$ ,  $1 \leq a \leq m$ , represent the equivalent classes of the reflected chain  $X_t^\mu$  with at least two elements, and  $H_a: \mathcal{D}_a \rightarrow \mathbb{R}$  the function (unique up to an additive constant) which turns the measure  $\mu$  conditioned to  $\mathcal{D}_a$  stationary for the chain induced by  $\mathfrak{M}_{H_a} \mathfrak{R}_{\mathcal{D}_a} \mathcal{L}$  (the generator  $\mathfrak{R}_{\mathcal{D}_a} \mathcal{L}$  tilted by  $H_a$ ).

*Proof.* By Lemma A.9,  $I_{\mathcal{L}}(\mu)$  is the sum of two terms. Consider  $\mu(V_0) I_{\mathcal{L}_0}(\mu_{V_0})$ . By Lemma A.7,  $I_{\mathcal{L}_0}(\mu_{V_0}) = K_{\mathcal{L}_0}(\mu_{V_0})$ , where  $K_{\mathcal{L}_0}$  is given by equation (A.15) with the set  $V$  replaced by  $V_0$  in the second sum. Clearly, the equivalent classes of the reflected chain  $X_t^\mu$  with at least two elements for the generator  $\mathcal{L}_0$  coincide with the ones for the generator  $\mathcal{L}$ . On the other hand there is no harm to replace in the second sum of (A.15) the condition  $x \in V_0$  by  $x \in V$  as  $R(x,y) = 0$  for all  $x \in V \setminus V_0$ . Hence,

$$\mu(V_0) I_{\mathcal{L}_0}(\mu_{V_0}) = - \sum_{a=1}^m \sum_{x \in \mathcal{D}_a} \sum_{y \in \mathcal{D}_a \setminus \{x\}} \mu(x) R_{H_a}(x,y) + \sum_{x \in V} \sum_{y \in V_0 \setminus \{x\}} \mu(x) R(x,y) .$$

To complete the proof of the corollary, it remains to recall the formula for  $I_{\mathcal{L}}(\mu)$  presented in Lemma A.9.  $\square$

## APPENDIX B. CONVERGENCE OF LEVEL 2 RATE FUNCTIONALS

In this section, we present some general results on the convergence of level 2 large deviations rate functionals. The first result asserts that the rate functionals converge provided the jump rates converge.

Denote by  $\mathcal{L}_n$ ,  $n \geq 1$ , the generator of a  $V$ -valued continuous-time Markov chain whose jump rates are represented by  $R_n(\cdot, \cdot)$ . Let  $I_n: \mathcal{P}(V) \rightarrow \mathbb{R}_+$  be the occupation time large deviations rate functional associated to the generator  $\mathcal{L}_n$ . This is the functional defined by formula (A.3) with the rates  $R_n$  replacing  $R$ . We first consider the case where  $\mathcal{L}$  is irreducible.

**Lemma B.1.** *Suppose that the Markov chains induced by the generators  $\mathcal{L}_n$ ,  $n \geq 1$ , and  $\mathcal{L}$  are irreducible, and that  $R_n(x, y) \rightarrow R(x, y) \in \mathbb{R}_+$  for all  $y \neq x \in V$ . Then,  $I_n(\mu) \rightarrow I(\mu)$  for all  $\mu \in \mathcal{P}(V)$  such that  $\mu(x) > 0$  for all  $x \in V$ . Here,  $I$  represents the rate functional associated to the jump rates  $R$ .*

*Proof.* Fix  $H: V \rightarrow \mathbb{R}$  and denote by  $J_H^{(n)}: \mathcal{P}(V) \rightarrow \mathbb{R}$  the functional given by (A.2) with  $R$  replaced by  $R_n$ . Then, as  $R_n(x, y) \rightarrow R(x, y)$ , for all  $H: V \rightarrow \mathbb{R}$

$$I_n(\mu) \geq J_H^{(n)}(\mu) \rightarrow J_H(\mu).$$

Maximizing over  $H$  yields that  $\liminf_{n \rightarrow \infty} I_n(\mu) \geq I(\mu)$ . Note that we did not use the irreducibility of  $\mathcal{L}$  in this part of the proof.

Conversely, since the Markov chain induced by the generator  $\mathcal{L}_n$  is irreducible and the support of  $\mu$  is the set  $V$ , by Lemmata A.3 and A.2,  $I_n(\mu) = J_{H_n}^{(n)}(\mu)$ , where  $H_n$  is the function which turns  $\mu$  the stationary stated for the tilted generator  $\mathfrak{M}_{H_n} \mathcal{L}_n$ .

Denote by  $\mathcal{E}$  the oriented edges  $(x, y) \in V \times V$  such that  $R(x, y) > 0$ . Let  $a = \min\{\mu(x) R(x, y) : (x, y) \in \mathcal{E}\} > 0$ . As  $R_n$  converges to  $R$ , there exists  $n_0 > 0$  such that  $\min\{\mu(x) R_n(x, y) : (x, y) \in \mathcal{E}\} \geq a/2$ . Since the Markov chain induced by the rates  $R(x, y)$  is irreducible and the rates  $R_n$  converge to  $R$ , by (A.8) and Remark A.4, there exists a finite constant  $C_0$  such that

$$\max_{y, x \in V} |H_n(y) - H_n(x)| \leq C_0 \tag{B.1}$$

for all  $n \geq n_0$ .

Therefore, there exist functions  $G: V \rightarrow \mathbb{R}$  and a subsequence  $n'$  such that  $H_{n'}(y) - H_{n'}(x) \rightarrow G(y) - G(x)$  for all  $x, y \in V$ . Hence, through this subsequence  $J_{H_n}^{(n)}(\mu)$  converges to  $J_H(\mu) \leq I(\mu)$ . This proves that  $\limsup_n I_n(\mu) \leq I(\mu)$  and completes the proof of the lemma.  $\square$

We now remove the assumption that  $\mathcal{L}$  is irreducible.

**Lemma B.2.** *Suppose that the Markov chain induced by the generator  $\mathcal{L}_n$  is irreducible for all  $n \geq 1$  and that  $R_n(x, y) \rightarrow R(x, y) \in \mathbb{R}_+$  for all  $y \neq x \in V$ . Then,  $I_n(\mu) \rightarrow I(\mu)$  for all  $\mu \in \mathcal{P}(V)$  such that  $\mu(x) > 0$  for all  $x \in V$ .*

*Proof.* In Lemma B.1, we proved that  $\liminf_{n \rightarrow \infty} I_n(\mu) \geq I(\mu)$ . Conversely, since the Markov chain induced by the generator  $\mathcal{L}_n$  is irreducible and the support of  $\mu$  is the set  $V$ , by Lemmata A.3 and A.2,  $I_n(\mu) = J_{H_n}^{(n)}(\mu)$ , where  $H_n$  is the function which turns  $\mu$  the stationary stated for the tilted generator  $\mathfrak{M}_{H_n} \mathcal{L}_n$ .

Denote by  $\mathfrak{Q}_a$ ,  $1 \leq a \leq \ell_\mu$  the equivalent classes of the generator  $\mathcal{L}$ , and by  $\mathfrak{D}_a$ ,  $1 \leq a \leq m_\mu$  the ones with at least two elements. By definition of  $J_{H_n}^{(n)}(\mu)$ ,

$$J_{H_n}^{(n)}(\mu) \leq \sum_{a=1}^{m_\mu} \mu(\mathfrak{D}_a) J_{\mathfrak{D}_a, H_n}^{(n)}(\mu_{\mathfrak{D}_a}) + \sum_{a=1}^{\ell_\mu} \sum_{b \neq a} \sum_{x \in \mathfrak{Q}_a} \sum_{y \in \mathfrak{Q}_b} \mu(x) R_n(x, y),$$

where  $J_{\mathfrak{D}_a, H_n}^{(n)}(\nu) = \sum_{x \in \mathfrak{D}_a} \sum_{y \in \mathfrak{D}_a \setminus \{x\}} \nu(x) R_n(x, y) [1 - e^{H_n(y) - H_n(x)}]$

As  $n \rightarrow \infty$ , the second term converges to the same sum with  $R$  in place of  $R_n$ . We turn to the first term.

Fix  $1 \leq a \leq m_\mu$ . By the arguments presented in the proof of Lemma B.1, (B.1) holds provided the maximum is carried over  $x, y \in \mathfrak{D}_a$ . Therefore, by the end of the proof of that lemma,  $\limsup_n J_{\mathfrak{D}_a, H_n}^{(n)}(\mu_{\mathfrak{D}_a}) \leq I_{\mathfrak{G}_{\mathfrak{D}_a} \mathcal{L}}(\mu_{\mathfrak{D}_a})$ . Recollecting the previous estimates and recalling Lemma A.7 and definition (A.14) yields that  $\limsup_n I_n(\mu) \leq I(\mu)$ . This completes the proof of the lemma.  $\square$

**$\Gamma$ -convergence.** In this subsection we present a result on  $\Gamma$ -convergence used in the article. Recall from Section 2 the definition of  $\Gamma$ -convergence. Fix a Polish space  $\mathcal{X}$  and a functional  $U: \mathcal{X} \rightarrow [0, +\infty]$ .

**Definition B.3.** A subset  $\mathcal{X}_0$  of  $\mathcal{X}$  is said to be  $U$ -dense if for every  $x \in \mathcal{X}$  such that  $U(x) < \infty$ , there exists a sequence  $(x_k : k \geq 1)$  such that  $x_k \in \mathcal{X}_0$ ,  $x_k \rightarrow x$  and  $U(x_k) \rightarrow U(x)$ .

**Lemma B.4.** Let  $\mathcal{X}_0$  be a  $U$ -dense subset of  $\mathcal{X}$ . To show that  $U$  is a  $\Gamma$ -limsup for the sequence  $U_n$ , it is enough to show that for every  $x \in \mathcal{X}_0$ , there exists a sequence  $(x_n : n \geq 1)$  such that  $x_n \rightarrow x$  and (2.2) holds.

*Proof.* Assume that for each  $x \in \mathcal{X}_0$ , there exists a sequence  $(x_n : n \geq 1)$  such that  $x_n \rightarrow x$  and (2.2) holds.

Fix  $x \in \mathcal{X}$ . We have to show that (2.2) holds for some sequence  $x_n \in \mathcal{X}$  which converges to  $x$ . If  $U(x) = \infty$ , there is nothing to prove. Assume, therefore, that  $U(x) < \infty$ . As  $\mathcal{X}_0$  is  $U$ -dense, there exists a sequence  $(x^{(k)} : k \geq 1)$  such that  $x^{(k)} \in \mathcal{X}_0$ ,  $x^{(k)} \rightarrow x$ , and  $U(x^{(k)}) \rightarrow U(x)$ .

Since the result holds for elements of  $\mathcal{X}_0$ , for each  $k \geq 1$ , there exists a sequence  $(x_n^{(k)} : n \geq 1)$  such that  $x_n^{(k)} \rightarrow x^{(k)}$ ,  $\limsup_{n \rightarrow \infty} U_n(x_n^{(k)}) \leq U(x^{(k)})$ . At this point, a classical diagonal argument permits to construct a sequence  $x_n$  such that  $x_n \rightarrow x$ ,  $\limsup_{n \rightarrow \infty} U_n(x_n) \leq U(x)$ , as claimed.  $\square$

We return to the context of the article and assume that  $\mathcal{X} = \mathcal{P}(V)$ .

**Lemma B.5.** Let  $\mathcal{P}_+$  be the subset of  $\mathcal{P}(V)$  formed by the measures whose support is  $V$ :  $\mathcal{P}_+ = \{\mu \in \mathcal{P}(V) : \mu(x) > 0 \forall x \in V\}$ . The set  $\mathcal{P}_+$  is  $I$ -dense.

*Proof.* Fix  $\mu \in \mathcal{P}(V)$  and let  $\nu$  be the uniform probability measure on  $V$ . Set  $\mu_n = [1 - (1/n)]\mu + (1/n)\nu$ . Clearly,  $\mu_n \in \mathcal{P}(V)$  and  $\mu_n \rightarrow \mu$ . It remains to show that  $I(\mu_n) \rightarrow I(\mu)$ .

Recall from Lemma A.1 the properties of the functional  $I$ . By the lower-semicontinuity of  $I$ ,  $I(\mu) \leq \liminf_n I(\mu_n)$ . By convexity,  $I(\mu_n) \leq [1 - (1/n)]I(\mu) + (1/n)I(\nu)$ . Since  $I(\nu) \leq |V|^{-1} \sum_{x \in V} \lambda(x) < \infty$ ,  $\limsup_n I(\mu_n) \leq I(\mu)$ , as claimed.  $\square$

## APPENDIX C. TRACE PROCESS AND LEVEL 2 RATE FUNCTIONALS

We examine in this section the effect of reducing the state space, by taking the trace of the process, on the large deviations rate functional. We first recall the definition of the trace process and some of its properties.

Denote by  $T^{\mathcal{A}}(t)$ ,  $\mathcal{A} \subsetneq V$ , the total time the process  $X_t$  spends in  $\mathcal{A}$  in the time-interval  $[0, t]$ :

$$T^{\mathcal{A}}(t) = \int_0^t \chi_{\mathcal{A}}(X_s) ds ,$$

where, recall,  $\chi_{\mathcal{A}}$  represents the indicator function of the set  $\mathcal{A}$ . Denote by  $S^{\mathcal{A}}(t)$  the generalized inverse of  $T^{\mathcal{A}}(t)$ :

$$S^{\mathcal{A}}(t) = \sup\{s \geq 0 : T^{\mathcal{A}}(s) \leq t\} .$$

The trace of  $X_t$  on  $\mathcal{A}$ , denoted by  $(X_t^{\mathcal{A}} : t \geq 0)$ , is defined by

$$X_t^{\mathcal{A}} = X_{S^{\mathcal{A}}(t)} ; \quad t \geq 0 . \quad (\text{C.1})$$

By Propositions 6.1 and 6.3 in [2], the trace process is an irreducible,  $\mathcal{A}$ -valued continuous-time Markov chain, obtained by turning off the clock when the process  $X_t$  visits the set  $\mathcal{A}^c$ , that is, by deleting all excursions to  $\mathcal{A}^c$ . For this reason, it is called the trace process of  $X_t$  on  $\mathcal{A}$ . For a  $V$ -valued Markov chain generator  $\mathcal{L}$  and a proper subset  $\mathcal{A}$  of  $V$ , denote by  $\mathfrak{T}_{\mathcal{A}}\mathcal{L}$  the generator of the trace process on  $\mathcal{A}$ .

Denote by  $R_{\mathcal{A}}^T(\cdot, \cdot)$  the jump rates of the trace process. Suppose that  $\mathcal{A} = V \setminus \{z\}$ . The first equation after [2, Corollary 6.2] asserts that for all  $x, y \in \mathcal{A}$ ,

$$R_{V \setminus \{z\}}^T(x, y) = R(x, y) + R(x, z)p(z, y) , \quad (\text{C.2})$$

where  $p(\cdot, \cdot)$  represents the jump probability of the chain  $X_t$ . In particular,  $R_{V \setminus \{z\}}^T(x, y) \geq R(x, y)$ . Iterating this procedure yields that

$$R_{\mathcal{A}}^T(x, y) \geq R(x, y) \quad (\text{C.3})$$

for all  $\mathcal{A} \subset V$ ,  $x \neq y \in \mathcal{A}$ .

We turn to the rate functional. To simplify certain formulae, we represent the rate function as

$$I(\mu) = \sup_u - \int_V \frac{\mathcal{L}u}{u} d\mu ,$$

where the supremum is carried out over all strictly positive functions  $u: V \rightarrow (0, \infty)$ . In consequence, in this section,  $\mathfrak{M}_u\mathcal{L}$  stands for the tilted generator given by

$$(\mathfrak{M}_u\mathcal{L}f)(x) = \sum_{y \in V} R_u(x, y) [f(y) - f(x)] , \quad R_u(x, y) = \frac{1}{u(x)} R(x, y) u(y) . \quad (\text{C.4})$$

Denote by  $I_{\mathcal{A}}^T$  the large deviations rate functional associated to the trace generator  $\mathfrak{T}_{\mathcal{A}}\mathcal{L}$ :

$$I_{\mathcal{A}}^T(\mu) := \sup_u - \int_{\mathcal{A}} \frac{(\mathfrak{T}_{\mathcal{A}}\mathcal{L})u}{u} d\mu , \quad \mu \in \mathcal{P}(\mathcal{A}) ,$$

where the sup is carried over all functions  $u: \mathcal{A} \rightarrow (0, \infty)$ .

Fix  $u: \mathcal{A} \rightarrow (0, \infty)$ , and denote by  $\mathfrak{H}u = \mathfrak{H}_{\mathcal{L}}u$  the  $\mathcal{L}$ -harmonic extension of  $u$  to  $V$ , that is, the solution of the Poisson equation

$$\begin{cases} \mathcal{L}v = 0, & V \setminus \mathcal{A}, \\ v = u, & \mathcal{A}. \end{cases} \quad (\text{C.5})$$

Next proposition is the main result of this section.

**Proposition C.1.** *Assume that the Markov chain induced by the generator  $\mathcal{L}$  is irreducible. Fix  $\mu \in \mathcal{P}(V)$  and assume that its support, denoted by  $\mathcal{A}$ , is a proper subset of  $V$ . Then,*

$$I_{\mathcal{A}}^{\mathbb{T}}(\mu) \leq I(\mu).$$

Conversely, let  $u: \mathcal{A} \rightarrow (0, \infty)$  such that

$$I_{\mathcal{A}}^{\mathbb{T}}(\mu) = - \int_{\mathcal{A}} \frac{(\mathfrak{T}_{\mathcal{A}}\mathcal{L})u}{u} d\mu.$$

Denote by  $v = \mathfrak{H}_{\mathcal{A}}u$  the harmonic extension of  $u$  to  $V$  given by the solution of (C.5). Let  $\nu$  be the stationary state of the tilted generator  $\mathfrak{M}_{\nu}\mathcal{L}$ . Then,

$$I_{\mathcal{A}}^{\mathbb{T}}(\mu) = \frac{1}{\nu(\mathcal{A})} I(\nu).$$

**Harmonic extension.** The harmonic extension has a stochastic representation. Denote by  $H_{\mathcal{A}}, H_{\mathcal{A}}^+, \mathcal{A} \subset V$ , the hitting and return time of  $\mathcal{A}$ :

$$H_{\mathcal{A}} := \inf \{t > 0 : X_t \in \mathcal{A}\}, \quad H_{\mathcal{A}}^+ := \inf \{t > \tau_1 : X_t \in \mathcal{A}\}, \quad (\text{C.6})$$

where  $\tau_1$  represents the time of the first jump of the chain  $X_t$ :  $\tau_1 = \inf \{t > 0 : X_t \neq X_0\}$ . By the strong Markov property, the solution of the Poisson equation can be represented as

$$(\mathfrak{H}u)(x) = \mathbf{E}_x[u(X_{H_{\mathcal{A}}})], \quad x \in V. \quad (\text{C.7})$$

In particular,

$$\min_{y \in \mathcal{A}} u(y) \leq \min_{x \in V} (\mathfrak{H}u)(x) \leq \max_{x \in V} (\mathfrak{H}u)(x) \leq \max_{y \in \mathcal{A}} u(y).$$

Moreover, by [5, Lemma A.1],

$$[(\mathfrak{T}_{\mathcal{A}}\mathcal{L})u](x) = [\mathcal{L}(\mathfrak{H}u)](x), \quad x \in \mathcal{A}. \quad (\text{C.8})$$

**Lemma C.2.** *Fix a subset  $\mathcal{A} \subsetneq \mathcal{B} \subsetneq V$ , and a function  $u: \mathcal{A} \rightarrow \mathbb{R}$ . Then,*

$$(\mathfrak{H}_{\mathfrak{T}_{\mathcal{B}}\mathcal{L}}u)(x) = (\mathfrak{H}_{\mathcal{L}}u)(x), \quad x \in \mathcal{B}.$$

This result asserts that the  $\mathcal{L}$ -harmonic extension of  $u$  to  $V$  coincides on the set  $\mathcal{B}$  with the  $(\mathfrak{T}_{\mathcal{B}}\mathcal{L})$ -harmonic extension of  $u$  to  $\mathcal{B}$ . More precisely, denote by  $v = \mathfrak{H}_{\mathcal{L}}u$  the  $\mathcal{L}$ -harmonic extension of  $u$  to  $V$  and by  $v_{\mathcal{B}}: \mathcal{B} \rightarrow \mathbb{R}$  its restriction to  $\mathcal{B}$ , defined as  $v_{\mathcal{B}}(x) = v(x)$ ,  $x \in \mathcal{B}$ . Lemma C.2 states that the function  $v_{\mathcal{B}}$  is the  $(\mathfrak{T}_{\mathcal{B}}\mathcal{L})$ -harmonic extension of  $u$  to  $\mathcal{B}$ . In other words, that  $v_{\mathcal{B}}$  solves equation (C.5) with  $\mathcal{B}, \mathfrak{T}_{\mathcal{B}}\mathcal{L}$  replacing  $V, \mathcal{L}$ , respectively.

*Proof of Lemma C.2.* Denote by  $X_t^{\mathcal{B}}$  the trace of the Markov chain  $X_t$  on  $\mathcal{B}$  and by  $H_{\mathcal{A}}(X^{\mathcal{B}})$  its hitting time of the set  $\mathcal{A}$ . Clearly, starting from  $x \in \mathcal{B}$ ,  $X_{H_{\mathcal{A}}(X^{\mathcal{B}})}^{\mathcal{B}} = X_{H_{\mathcal{A}}}^{\mathcal{B}}$  almost surely, so that

$$\mathbf{E}_x[u(X_{H_{\mathcal{A}}(X^{\mathcal{B}})}^{\mathcal{B}})] = \mathbf{E}_x[u(X_{H_{\mathcal{A}}})]$$

for all  $x \in \mathcal{B}$ . By (C.7), the left-hand side of this equation is  $(\mathfrak{H}_{\mathfrak{T}_{\mathcal{B}}\mathcal{L}} u)(x)$ , and the right-hand side is  $(\mathfrak{H}_{\mathcal{L}} u)(x)$ . This completes the proof of the lemma.  $\square$

Enumerate the set  $V \setminus \mathcal{A}$  as  $\{x_1, \dots, x_p\}$ , and let  $\mathcal{A}_0 = V$ ,  $\mathcal{A}_k = V \setminus \{x_1, \dots, x_k\}$ ,  $1 \leq k \leq p$ , so that  $\mathcal{A} = \mathcal{A}_p$ . Fix  $u: \mathcal{A} \rightarrow (0, \infty)$ , and denote by  $v$  its  $\mathcal{L}$ -harmonic extension to  $V$ , given by (C.5). Let  $v_k: \mathcal{A}_k \rightarrow (0, \infty)$  be the restriction of  $v$  to  $\mathcal{A}_k$ ,  $1 \leq k \leq p$ , and  $v_0 = v$ .

**Corollary C.3.** *For all  $1 \leq k \leq p$ ,*

$$[(\mathfrak{T}_{\mathcal{A}_{k-1}}\mathcal{L})v_{k-1}](y) = 0, \quad y \in \mathcal{A}_{k-1} \setminus \mathcal{A}.$$

*In particular,  $v_{k-1}$  is the  $(\mathfrak{T}_{\mathcal{A}_{k-1}}\mathcal{L})$ -harmonic extension of  $v_j$  for  $k \leq j \leq p$ .*

*Proof.* Fix  $1 \leq k \leq p$ . By Lemma C.2,  $v_{k-1}$  is the  $(\mathfrak{T}_{\mathcal{A}_{k-1}}\mathcal{L})$ -harmonic extension of  $u$  to  $\mathcal{A}_{k-1}$ . Hence, for all  $y \in \mathcal{A}_{k-1} \setminus \mathcal{A}$ ,  $[(\mathfrak{T}_{\mathcal{A}_{k-1}}\mathcal{L})v_{k-1}](y) = 0$ , as claimed. By definition, for  $k \leq j \leq p$ ,  $v_{k-1}$  and  $v_j$  coincide on  $\mathcal{A}_j$ , which proves the second assertion of the corollary.  $\square$

**Tilted dynamics.** Fix  $v: V \rightarrow (0, \infty)$ , and recall the definition of the jump rates  $R_v$  introduced in (C.4). Denote by  $\lambda_v(x)$ ,  $\lambda(x)$  the holding times at  $x$  associated to the rates  $R_v(\cdot, \cdot)$ ,  $R(\cdot, \cdot)$ , respectively. Assume that  $(\mathcal{L}v)(x) = 0$ . Then,

$$\lambda_v(x) = \lambda(x) \tag{C.9}$$

Indeed, since  $(\mathcal{L}v)(x) = 0$ ,

$$\sum_{y \in V} R(x, y) v(y) = \sum_{y \in V} R(x, y) v(x).$$

Hence,

$$\lambda_v(x) = \frac{1}{v(x)} \sum_{y \in V} R(x, y) v(y) = \frac{1}{v(x)} \sum_{y \in V} R(x, y) v(x) = \lambda(x),$$

as claimed.

**Lemma C.4.** *Fix  $u: \mathcal{A} \rightarrow (0, \infty)$ . Let  $v: V \rightarrow (0, \infty)$  be its harmonic extension to  $V$  defined by (C.5). Then,*

$$\mathfrak{T}_{\mathcal{A}}(\mathfrak{M}_v\mathcal{L}) = \mathfrak{M}_u(\mathfrak{T}_{\mathcal{A}}\mathcal{L}).$$

*Proof.* We first prove the lemma for  $\mathcal{A} = V \setminus \{x_0\}$ . Recall that we denote by  $R_v(x, y)$  the jump rates of the tilted generator  $\mathfrak{M}_v\mathcal{L}$ . Denote by  $R_{v,\mathcal{A}}(x, y)$  the jump rates of the trace generator  $\mathfrak{T}_{\mathcal{A}}\mathfrak{M}_v\mathcal{L}$ . By (C.2), for all  $x \neq y \in \mathcal{A}$ ,

$$R_{v,\mathcal{A}}(x, y) = R_v(x, y) + \frac{1}{\lambda_v(x_0)} R_v(x, x_0) R_v(x_0, y),$$

where, recall,  $\lambda_v(x_0)$  stands for the holding time at  $x_0$  associated to the rates  $R_v(x, y)$ . Since  $v$  is harmonic at  $x_0$ , by (C.9) and the definition of  $R_v$ , the previous expression is equal to

$$\frac{v(y)}{v(x)} \left\{ R(x, y) + \frac{1}{\lambda(x_0)} R(x, x_0) R(x_0, y) \right\}.$$

Since  $v$  and  $u$  coincide on  $\mathcal{A}$ , we may replace  $v(x)$ ,  $v(y)$  by  $u(x)$ ,  $u(y)$ , respectively. Representing by  $R_{\mathcal{A}}(z, w)$  the jump rates associated to the generator  $\mathfrak{T}_{\mathcal{A}}\mathcal{L}$ , by (C.2) once more, the previous expression is equal to

$$\frac{u(y)}{u(x)} R_{\mathcal{A}}(x, y),$$

which proves the lemma in the case where  $V \setminus \mathcal{A}$  is a singleton.

We extend the result to arbitrary sets  $\mathcal{A}$ . Fix a set  $\mathcal{A}$ , and recall the notation introduced above Corollary C.3. Fix  $1 \leq j \leq p$ . By Corollary C.3,  $v_{k-1}$  is the  $(\mathfrak{T}_{\mathcal{A}_{k-1}}\mathcal{L})$ -harmonic extension of  $v_k$ . Thus, by the first part of the proof applied to the generator  $\mathfrak{T}_{\mathcal{A}_{k-1}}\mathcal{L}$ ,

$$\mathfrak{T}_{\mathcal{A}_k} \mathfrak{M}_{v_{k-1}} \mathfrak{T}_{\mathcal{A}_{k-1}} \mathcal{L} = \mathfrak{M}_{v_k} \mathfrak{T}_{\mathcal{A}_k} \mathfrak{T}_{\mathcal{A}_{k-1}} \mathcal{L} = \mathfrak{M}_{v_k} \mathfrak{T}_{\mathcal{A}_k} \mathcal{L} \quad (\text{C.10})$$

because  $\mathfrak{T}_{\mathcal{A}_k} \mathfrak{T}_{\mathcal{A}_{k-1}} \mathcal{L} = \mathfrak{T}_{\mathcal{A}_k} \mathcal{L}$ .

Since  $\mathcal{A}_0 = V$ ,  $\mathfrak{T}_{\mathcal{A}_0} \mathcal{L} = \mathcal{L}$ , and (C.10) for  $k = 1$  states that

$$\mathfrak{T}_{\mathcal{A}_1} \mathfrak{M}_{v_0} \mathcal{L} = \mathfrak{M}_{v_1} \mathfrak{T}_{\mathcal{A}_1} \mathcal{L}.$$

Applying  $\mathfrak{T}_{\mathcal{A}_2}$  on both sides of this identity and then (C.10) yields that

$$\mathfrak{T}_{\mathcal{A}_2} \mathfrak{M}_{v_0} \mathcal{L} = \mathfrak{M}_{v_2} \mathfrak{T}_{\mathcal{A}_2} \mathcal{L}.$$

Iterating this procedure completes the proof of the lemma as  $v_0 = v$ ,  $v_p = u$ ,  $\mathcal{A}_0 = V$ ,  $\mathcal{A}_p = \mathcal{A}$ .  $\square$

**Proof of Proposition C.1.** By [2, Proposition 6.1], the Markov chain induced by the trace generator  $\mathfrak{T}_{\mathcal{A}}\mathcal{L}$  is irreducible. Hence, since  $\mu(x) > 0$  for all  $x \in \mathcal{A}$ , by Lemma A.3, there exists  $u: \mathcal{A} \rightarrow (0, \infty)$  such that

$$I_{\mathcal{A}}^{\Gamma}(\mu) = - \int_{\mathcal{A}} \frac{(\mathfrak{T}_{\mathcal{A}}\mathcal{L})u}{u} d\mu.$$

Denote by  $v$  the harmonic extension of  $u$  to  $V$  given by the solution of (C.5). By (C.8),  $(\mathcal{L}v)(x) = [(\mathfrak{T}_{\mathcal{A}}\mathcal{L})u](x)$  for all  $x \in \mathcal{A}$ . Hence, the right-hand side of the previous displayed equation is equal to

$$- \int_{\mathcal{A}} \frac{\mathcal{L}v}{v} d\mu = - \int_V \frac{\mathcal{L}v}{v} d\mu \leq I(\mu).$$

The identity follows from the fact that  $\mathcal{A}$  is the support of  $\mu$  (or from the fact that  $v$  is harmonic in  $\mathcal{A}^c$ ), and the inequality from the definition of the functional  $I$ .

We turn to the converse assertion. By definition of  $u$ ,  $v$  and [5, Lemma A.1],

$$I_{\mathcal{A}}^{\Gamma}(\mu) = - \int_{\mathcal{A}} \frac{(\mathfrak{T}_{\mathcal{A}}\mathcal{L})u}{u} d\mu = - \int_{\mathcal{A}} \frac{\mathcal{L}v}{v} d\mu.$$

Recall that  $\nu$  represents the stationary state of the tilted generator  $\mathfrak{M}_v\mathcal{L}$ . By [2, Proposition 6.3], the measure  $\nu$  conditioned to  $\mathcal{A}$  is the stationary state of the trace (the Markov chain induced by the generator  $\mathfrak{T}_{\mathcal{A}}\mathfrak{M}_v\mathcal{L}$ ). Since, by Lemma C.4,  $\mathfrak{T}_{\mathcal{A}}\mathfrak{M}_v\mathcal{L} = \mathfrak{M}_u\mathfrak{T}_{\mathcal{A}}\mathcal{L}$ ,  $\nu$  conditioned to  $\mathcal{A}$  is the stationary state of the chain associated to  $\mathfrak{M}_u\mathfrak{T}_{\mathcal{A}}\mathcal{L}$ . By Lemma A.1,  $\mu$  is stationary for  $\mathfrak{M}_u\mathfrak{T}_{\mathcal{A}}\mathcal{L}$  as well. Since the chain  $X_t$  is irreducible, so is the trace and the tilted trace. Thus, by uniqueness,  $\mu(\cdot) = \nu(\cdot)/\nu(\mathcal{A})$ , and the right-hand side of the previous displayed equation can be written as

$$- \frac{1}{\nu(\mathcal{A})} \int_{\mathcal{A}} \frac{\mathcal{L}v}{v} d\nu = - \frac{1}{\nu(\mathcal{A})} \int_V \frac{\mathcal{L}v}{v} d\nu$$

because  $v$  is harmonic on  $V \setminus \mathcal{A}$ . Since  $\nu$  is stationary for the tilted generator  $\mathfrak{M}_v \mathcal{L}$ , by Lemma A.2, the right-hand side is equal to  $\nu(\mathcal{A})^{-1}I(\nu)$ . This completes the proof of the lemma.  $\square$

**Acknowledgments.** C. L. has been partially supported by FAPERJ CNE E-26/201.207/2014 and by CNPq Bolsa de Produtividade em Pesquisa PQ 303538/2014-7.

## REFERENCES

- [1] I. Armendáriz, S. Großkinsky, M. Loulakis. Zero range condensation at criticality. *Stochastic Process. Appl.* **123**, 346–3496 (2013).
- [2] J. Beltrán, C. Landim: Tunneling and metastability of continuous time Markov chains. *J. Stat. Phys.* **140** 1065–1114 (2010).
- [3] J. Beltrán, C. Landim: Metastability of reversible condensed zero range processes on a finite set, *Probab. Theory Relat. Fields* **152**, 781–807 (2012).
- [4] J. Beltrán, C. Landim; Tunneling and metastability of continuous time Markov chains II. *J. Stat. Phys.* **149**, 598–618 (2012).
- [5] L. Bertini, D. Gabrielli, C. Landim: Metastable  $\Gamma$ -expansion of finite state Markov chains level two large deviations rate functions. arXiv:2207.02588, (2022).
- [6] A. Bianchi, S. Dommers, C. Giardinà: Metastability in the reversible inclusion process. *Electron. J. Probab.* **22**, paper no. 70, (2017).
- [7] A. Bovier, M. Eckhoff, V. Gayraud, M. Klein: Metastability in reversible diffusion process I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc.* **6**, 399–424 (2004)
- [8] A. Braides;  *$\Gamma$ -Convergence for beginners*. Oxford University Press, Oxford, 2002.
- [9] J. Cao, P. Chleboun, S. Grosskinsky: Dynamics of condensation in the totally asymmetric inclusion process. *J. Stat. Phys.* **155**, 523–543 (2014).
- [10] M. D. Donsker, S. R. S. Varadhan: Asymptotic evaluation of certain Markov proces expectations for large time, I, *Comm. Pure Appl. Math.*, **28**, 1-47 (1975).
- [11] M. Freidlin, L. Koralov: Metastable Distributions of Markov Chains with Rare Transitions. *J. Stat. Phys.* **167**, 1355–1375 (2017)
- [12] G. Di Gesù, M. Mariani: Full metastable asymptotic of the Fisher information. *SIAM J. MATH. ANAL.* **49**, 3048–3072 (2017)
- [13] S. Grosskinsky, F. Redig, K. Vafayi: Dynamics of condensation in the symmetric inclusion process **18**, article no. 66 (2013).
- [14] S. Kim. Second time scale of the metastability of reversible inclusion processes. *Probab. Theory Relat. Fields* **180**, 1135–1187 (2021).
- [15] S. Kim, I. Seo: Condensation and Metastable Behavior of Non-reversible Inclusion Processes. *Commun. Math. Phys.* **382**, 1343–1401 (2021).
- [16] C. Landim: Metastability for a non-reversible dynamics: the evolution of the condensate in totally asymmetric zero range processes. *Commun. Math. Phys.* **330**, 1–32 (2014).
- [17] C. Landim: Metastable Markov chains. *Probability Surveys* **16**, 143–227 (2019). DOI: 10.1214/18-PS310
- [18] C. Landim, M. Loulakis, M. Mourragui: Metastable Markov chains: from the convergence of the trace to the convergence of the finite-dimensional distributions. *Electron. J. Probab.* **23**, paper no. 95 (2018).
- [19] C. Landim, M. Mariani, I. Seo: A Dirichlet and a Thomson principle for non-selfadjoint elliptic operators, *Metastability in non-reversible diffusion processes*. *Arch. Rational Mech. Anal.* **231**, 887–938 (2019)
- [20] C. Landim, R. Misturini, K. Tsunoda: Metastability of reversible random walks in potential field. *J. Stat. Phys.* **160**, 1449–1482 (2015)
- [21] C. Landim, I. Seo: Metastability of non-reversible random walks in a potential field, the Eyring-Kramers transition rate formula. *Comm. Pure. Appl. Math.* **71**, 203–266 (2018)
- [22] C. Landim, I. Seo: Metastability of one-dimensional, non-reversible diffusions with periodic boundary conditions. *Ann. Inst. H. Poincaré, Probab. Statist.* **55**, 1850–1889 (2019).
- [23] C. Landim, T. Xu; Metastability of finite state Markov chains: a recursive procedure to identify slow variables for model reduction. *ALEA Lat. Am. J. Probab. Math. Stat.* **13**, 725-751 (2016)



- [24] J. Lee, I. Seo: Non-reversible metastable diffusions with Gibbs invariant measure I: Eyring–Kramers formula. *Probab. Theory Relat. Fields* **182**, 849–903 (2022).
- [25] J. Lee, I. Seo: Non-reversible metastable diffusions with Gibbs invariant measure II: Markov chain convergence. *arXiv:2008.08295* (2022).
- [26] M. Mariani: A  $\Gamma$ -convergence approach to large deviations, *arXiv:1204.0640* (2012).
- [27] C. Oh, F. Rezakhanlou: Metastability of zero range processes via Poisson equations. preprint available at [math.berkeley.edu](http://math.berkeley.edu) (2019).
- [28] F. Rezakhanlou, I. Seo: Scaling limit of small random perturbation of dynamical systems *arXiv:1812.02069* (2018).
- [29] I. Seo: Condensation of non-reversible zero-range processes, *Commun. Math. Phys.* **366**, 781–839 (2019)
- [30] M. Sugiura: Metastable behaviors of diffusion processes with small parameter. *J. Math. Soc. Japan* **47**, 755–788 (1995)

CLAUDIO LANDIM

IMPA

ESTRADA DONA CASTORINA 110,

J. BOTANICO, 22460 RIO DE JANEIRO, BRAZIL

and

CNRS UMR 6085, UNIVERSITÉ DE ROUEN,

AVENUE DE L'UNIVERSITÉ, BP.12, TECHNOPOLE DU MADRILLET,

F76801 SAINT-ÉTIENNE-DU-ROUVRAY, FRANCE.

*Email address:* `landim@impa.br`