ALEA, Lat. Am. J. Probab. Math. Stat. **13**, 725–751 (2016)
DOI: 10.30757/ALEA.v13-29

Metastability of finite state Markov chains: a recursive procedure to identify slow variables for model reduction

Claudio Landim and Tiecheng Xu

IMPA, Estrada Dona Castorina 110, CEP 22460-320 Rio de Janeiro, Brasil.
CNRS UMR 6085, Université de Rouen, Avenue de l'Université, BP.12, Technopôle du Madrillet, F76801 Saint-Étienne-du-Rouvray, France.
E-mail address: landim@impa.br

IMPA, Estrada Dona Castorina 110, CEP 22460-320 Rio de Janeiro, Brasil.
E-mail address: tcxu@impa.br

Abstract. Consider a sequence $(\eta^N(t) : t \geq 0)$ of continuous-time, irreducible Markov chains evolving on a fixed finite set E. Denote by $R_N(\eta, \xi)$ the jump rates of the Markov chain η_t^N, and assume that for any pair of bonds (η, ξ), (η', ξ') $\arctan\{R_N(\eta, \xi)/R_N(\eta', \xi')\}$ converges as $N \uparrow \infty$. Under a hypothesis slightly more restrictive (cf. (2.6) below), we present a recursive procedure which provides a sequence of increasing time-scales $\theta_N^1, \dots, \theta_N^{\mathfrak{p}}$, $\theta_N^j \ll \theta_N^{j+1}$, and of coarsening partitions of the set E, $\{\mathcal{E}_1^j, \dots, \mathcal{E}_{\mathfrak{n}_j}^j, \Delta^j\}$, $1 \leq j \leq \mathfrak{p}$, with the following property. Let $\phi_j : E \to \{0, 1, \dots, \mathfrak{n}_j\}$ be the projection defined by $\phi_j(\eta) = \sum_{x=1}^{\mathfrak{n}_j} x\, \mathbf{1}\{\eta \in \mathcal{E}_x^j\}$. For each $1 \leq j \leq \mathfrak{p}$, we prove that the hidden Markov chain $X_N^j(t) = \phi_j(\eta^N(t\theta_N^j))$ converges to a Markov chain on $\{1, \dots, \mathfrak{n}_j\}$.

1. Introduction

This article has two motivations. On the one hand, the metastable behavior of non-reversible Markovian dynamics has attracted much attention recently, see Metzner et al. (2009); Maes and O'Kelly de Galway (2013); Benois et al. (2013); Landim (2014); Misturini (2016); Cirillo et al. (2015); Bianchi and Gaudillière (2016); Chleboun and Grosskinsky (2015); Fernandez et al. (2015, 2016). On the other hand, the emergence of large complex networks gives a particular importance to the problem of data and model reduction, see Lu and Vanden-Eijnden (2014); Cameron and Vanden-Eijnden (2014); Avena and Gaudillière (2013). This issue arises in as diverse contexts as meteorology, genetic networks or protein folding, and is very

Received by the editors December 23, 2015; accepted August 1st, 2016.
2010 *Mathematics Subject Classification.* 60J27, 60J28.
Key words and phrases. Metastability, Markov chains, slow variables, model reduction.

closely related to the identification of slow variables, a fundamental tool in decreasing the degrees of freedom of a system, see Singer et al. (2009).

Not long ago, Beltrán and one of the authors of this article introduced a general approach to derive the metastable behavior of continuous-time Markov chains, particularly convenient in the presence of several valleys with the same depth, see Beltrán and Landim (2010, 2012, 2015). In the context of finite state Markov chains, see Beltrán and Landim (2011), it permits to identify the slow variables and to reduce the model.

More precisely, denote by E a finite set, by η_t^N a sequence of E-valued continuous-time, irreducible Markov chains, and by $\mathcal{E}_1, \dots, \mathcal{E}_{\mathfrak{n}}, \Delta$ a partition of the set E. Let $\mathcal{E} = \cup_{1 \le x \le \mathfrak{n}} \mathcal{E}_x$ and let $\phi_{\mathcal{E}} : E \to \{0, 1, \dots, \mathfrak{n}\}$ be the projection defined by

$$\phi_{\mathcal{E}}(\eta) \;=\; \sum_{x=1}^{\mathfrak{n}} x\, \mathbf{1}\{\eta \in \mathcal{E}_x\}\,.$$

In general, $X_N(t) = \phi_{\mathcal{E}}(\eta_t^N)$ is not a Markov chain, but only a hidden Markov chain. We say that $\phi_{\mathcal{E}}$ is a slow variable if there exists a time-scale θ_N for which the dynamics of $X_N(t\theta_N)$ is asymptotically Markovian.

The set Δ plays a special role in the partition, separating the sets $\mathcal{E}_1, \dots, \mathcal{E}_{\mathfrak{n}}$, called valleys. The chain remains a negligible amount of time in the set Δ in the time-scale θ_N at which the slow variable evolves.

Slow variables provide an efficient mechanism to contract the state space and to reduce the model in complex networks, as it allows to represent the original evolution through a simple Markovian chain $X_N(t)$ which takes value in a much smaller set, without losing the essential features of the dynamics. It may also reveal aspects of the dynamics which may not be apparent at first sight.

When the number of sets in the partition is reduced to 2, $\mathfrak{n} = 2$, and the Markov chain which describes the asymptotic behavior of the slow variable has one absorbing point and one transient point, the chain presents a metastable behavior. In a certain time-scale, it remains for an exponential time on a subset of the state space after which it jumps to another set where it remains for ever. By extension, and may be inappropriately, we say that the chain η_t^N exhibits a metastable behavior among the valleys $\mathcal{E}_1, \dots, \mathcal{E}_{\mathfrak{n}}$ in the time-scale θ_N whenever we prove the existence of a slow variable.

We present in this article a recursive procedure which permits to determine all slow variables of the chain. It provides a sequence of time-scales $\theta_N^1, \dots, \theta_N^{\mathfrak{p}}$ and of partitions $\{\mathcal{E}_1^j, \dots, \mathcal{E}_{\mathfrak{n}_j}^j, \Delta_j\}$, $1 \le j \le \mathfrak{p}$, of the set E with the following properties.

- The time-scales are increasing: $\lim_{N\to\infty} \theta_N^j / \theta_N^{j+1} = 0$ for $1 \le j < \mathfrak{p}$. This relation is represented as $\theta_N^j \ll \theta_N^{j+1}$.
- The partitions are coarser. Each set of the $(j+1)$-th partition is obtained as a union of sets in the j-th partition. Thus $\mathfrak{n}_{j+1} < \mathfrak{n}_j$ and for each a in $\{1, \dots, \mathfrak{n}_{j+1}\}$, $\mathcal{E}_a^{j+1} = \cup_{x \in A} \mathcal{E}_x^j$ for some subset A of $\{1, \dots, \mathfrak{n}_j\}$.
- The sets Δ_j, which separates the valleys, increase: $\Delta_j \subset \Delta_{j+1}$. Actually, $\Delta_{j+1} = \Delta_j \cup_{x \in B} \mathcal{E}_x^j$ for some subset B of $\{1, \dots, \mathfrak{n}_j\}$.
- The projection $\Psi_N^j(\eta) \;=\; \sum_{1 \le x \le \mathfrak{n}_j} x\, \mathbf{1}\{\eta \in \mathcal{E}_x^j\} \;+\; N\, \mathbf{1}\{\eta \in \Delta_j\}$ is a slow variable which evolves in the time-scale θ_N^j.

We prove three further properties of the partitions $\{\mathcal{E}_1^j, \dots, \mathcal{E}_{\mathfrak{n}_j}^j, \Delta_j\}$.

- As mentioned above, the amount of time the chain remains in the set Δ_j in the time-scale θ_N^j is negligible. We refer to condition (H3) below for a mathematical formulation of this assertion.
- Starting from any configuration in $\mathcal{E}_x^j$, the chain η_t^N attains the set $\cup_{y\neq x}\mathcal{E}_y^j$ at a time which is asymptotically exponential in the time-scale θ_N^j (cf. Remark 2.4).
- With a probability asymptotically equal to 1, the chain η_t^N visits all points of the set $\mathcal{E}_x^j$ before hitting another set $\mathcal{E}_y^j$ of the partition. In the terminology of Freidlin and Wentzell (1998), the sets of the first partition, denoted by $\mathcal{E}_x^1$, are cycles while the set of the following partitions are cycles of cycles.

These results have been proved in Beltrán and Landim (2011) for finite state reversible Markovian dynamics. We remove in this article the assumption of reversibility and we simplify some proofs.

In contrast with other approaches, see Manzo et al. (2004); Olivieri and Vares (2005); E and Vanden-Eijnden (2006); Metzner et al. (2009); Cirillo et al. (2015); Fernandez et al. (2015, 2016), we do not describe the tube of typical trajectories in a transition between two valleys, nor do we identify the critical configurations which are visited with high probability in such transitions.

The arguments presented here have been designed for sequences of Markov chains. The examples we have in mind are zero-temperature limits of non-reversible dynamics in a finite state space. It is not clear whether the analysis can be adapted to handle the case of a single fixed dynamics as in Cameron and Vanden-Eijnden (2014); Lu and Vanden-Eijnden (2014); Avena and Gaudillière (2013).

The approach presented in this article is based on a multiscale analysis. The sequence of increasing time-scales is defined in terms of the depth of the different valleys. In this sense, the method is similar to the one proposed by Scoppola (1993), and developed by Olivieri and Scoppola (1995, 1996), but it does not require the valleys to have exponential depth, nor the jump rates to be expressed in terms of exponentials. Actually, one of its main merit is that it relies on a minimal hypothesis, presented in (2.6) below, which is very easy to check since it is formulated only in terms of the jump rates.

The article is organized as follows. In Section 2 we state the main results. In the following three sections we introduce the tools needed to prove these results, which is carried out in the last three sections.

2. Notation and main results

This section is divided in four subsections. In the first one, we establish the notation and recall some known results which are used throughout the article. In the second subsection, we introduce the central hypothesis of the article, and, in the third one, we state the first main result of the article, Theorem 2.7, which describes the metastable behavior of the chain in the slowest time-scale. Finally, in the last subsection, we present in Theorem 2.12 all the time scales at which a metastable behavior is observed.

2.1. *Preliminaries.* Consider a finite set E. The elements of E are called configurations and are denoted by the Greek letters η, ξ, ζ. Consider a sequence of

continuous-time, E-valued, irreducible Markov chains $\{\eta_t^N : t \geq 0\}$. Denote the jump rates of η_t^N by $R_N(\eta, \xi)$, and by μ_N the unique invariant probability measure.

Denote by $D(\mathbb{R}_+, E)$ the space of right-continuous functions $x : \mathbb{R}_+ \to E$ with left-limits endowed with the Skorohod topology, and by $\mathbb{P}_\eta = \mathbb{P}_\eta^N$, $\eta \in E$, the probability measure on the path space $D(\mathbb{R}_+, E)$ induced by the Markov chain η_t^N starting from η. Expectation with respect to $\mathbb{P}_\eta$ is represented by $\mathbb{E}_\eta$.

Fix a proper subset A of E and denote by η_t^A the trace of the Markov chain η_t^N on the set A. The trace process is obtained by stopping the clock when the chain leaves the set A and by switching it on when it returns to the set A. More precisely, denote by $\boldsymbol{T}(t)$ the total time spent on A before time t:

$$\boldsymbol{T}(t) \;=\; \int_0^t \mathbf{1}\{\eta_s^N \in A\}\, ds\,,$$

where $\mathbf{1}\{B\}$ represents the indicator of the set B. Note that the function $\boldsymbol{T}$ is piecewise differentiable and that its derivative takes only the values 1 and 0. It is equal to 1 when the process is in A and it is equal to 0 when it is not. Let $\boldsymbol{S}(t)$ be the generalized inverse of $\boldsymbol{T}$:

$$\boldsymbol{S}(t) = \sup\{s \geq 0 : \boldsymbol{T}(s) \leq t\}\,.$$

The trace process is defined as $\eta_t^A = \eta^N_{\boldsymbol{S}(t)}$. It is shown in Beltrán and Landim (2010, Proposition 6.1) that η_t^A is a continuous-time, A-valued, irreducible Markov chain whose jump rates can be expressed in terms of the probabilities of hitting times of the original chain.

Denote by $H_{\mathcal{A}}$, $H_{\mathcal{A}}^+$, $\mathcal{A} \subset E$, the hitting time and the time of the first return to $\mathcal{A}$:

$$H_{\mathcal{A}} \;=\; \inf\big\{t > 0 : \eta_t^N \in \mathcal{A}\big\}\,, \quad H_{\mathcal{A}}^+ \;=\; \inf\big\{t > \tau_1 : \eta_t^N \in \mathcal{A}\big\}\,, \tag{2.1}$$

where τ_1 represents the time of the first jump of the chain η_t^N: $\tau_1 = \inf\{t > 0 : \eta_t^N \neq \eta_0^N\}$.

Denote by $\lambda_N(\eta)$, $\eta \in E$, the holding rates of the Markov chain η_t^N and by $p_N(\eta, \xi)$, η, $\xi \in E$, the jump probabilities, so that $R_N(\eta, \xi) = \lambda_N(\eta) p_N(\eta, \xi)$. For two disjoint subsets $\mathcal{A}$, $\mathcal{B}$ of E, denote by $\mathrm{cap}_N(\mathcal{A}, \mathcal{B})$ the capacity between $\mathcal{A}$ and $\mathcal{B}$:

$$\mathrm{cap}_N(\mathcal{A}, \mathcal{B}) \;=\; \sum_{\eta \in \mathcal{A}} \mu_N(\eta)\, \lambda_N(\eta)\, \mathbb{P}_\eta[H_{\mathcal{B}} < H_{\mathcal{A}}^+]\,. \tag{2.2}$$

Consider a partition $\mathcal{E}_1, \dots, \mathcal{E}_{\mathfrak{n}}, \Delta$ of the set E, which does not depend on the parameter N and such that $\mathfrak{n} \geq 2$. Fix two sequences of positive real numbers α_N, θ_N such that $\alpha_N \ll \theta_N$, where this notation stands for $\lim_{N\to\infty} \alpha_N/\theta_N = 0$.

Let $\mathcal{E} = \cup_{x \in S} \mathcal{E}_x$, where $S = \{1, \dots, \mathfrak{n}\}$. Denote by $\{\eta_t^{\mathcal{E}} : t \geq 0\}$ the trace of $\{\eta_t^N : t \geq 0\}$ on $\mathcal{E}$, and by $R_N^{\mathcal{E}} : \mathcal{E} \times \mathcal{E} \to \mathbb{R}_+$ the jump rates of the trace process $\eta_t^{\mathcal{E}}$:

$$R_N^{\mathcal{E}}(\eta, \xi) \;=\; \lambda_N(\eta)\, \mathbb{P}_\eta^N\big[H_\xi = H_{\mathcal{E}}^+\big]\,, \quad \eta\,,\ \xi \in \mathcal{E}\,,\ \eta \neq \xi\,.$$

We refer to Section 6 of Beltrán and Landim (2010) for a proof of this identity. Denote by $r_N^{\mathcal{E}}(\mathcal{E}_x, \mathcal{E}_y)$ the mean rate at which the trace process jumps from $\mathcal{E}_x$ to $\mathcal{E}_y$:

$$r_N^{\mathcal{E}}(\mathcal{E}_x, \mathcal{E}_y) \;=\; \frac{1}{\mu_N(\mathcal{E}_x)} \sum_{\eta \in \mathcal{E}_x} \mu_N(\eta) \sum_{\xi \in \mathcal{E}_y} R_N^{\mathcal{E}}(\eta, \xi)\,. \tag{2.3}$$

Assume that for every $x \neq y \in S$,

$$\begin{aligned} r_{\mathcal{E}}(x,y) \;&:=\; \lim_{N\to\infty} \theta_N\, r_N^{\mathcal{E}}(\mathcal{E}_x, \mathcal{E}_y) \;\in\; \mathbb{R}_+ \,, \\ \text{and that} \quad & \sum_{x\in S} \sum_{y\neq x} r_{\mathcal{E}}(x,y) > 0 \,. \end{aligned} \tag{H1}$$

The symbol := in the first line of the previous displayed equation means that the limit exists, that it is denoted by $r_{\mathcal{E}}(x,y)$, and that it belongs to $\mathbb{R}_+$. This convention is used throughout the article.

Assume that for every $x \in S$ for which $\mathcal{E}_x$ is not a singleton and for all $\eta \neq \xi \in \mathcal{E}_x$,

$$\liminf_{N\to\infty} \alpha_N \frac{\mathrm{cap}_N(\eta,\xi)}{\mu_N(\mathcal{E}_x)} \;>\; 0 \,. \tag{H2}$$

Finally, assume that in the time scale θ_N the chain remains a negligible amount of time outside the set $\mathcal{E}$: For every $t > 0$,

$$\lim_{N\to\infty} \max_{\eta\in E} \mathbb{E}_\eta\Big[\int_0^t \mathbf{1}\{\eta^N_{s\theta_N} \in \Delta\}\, ds \Big] \;=\; 0 \,. \tag{H3}$$

Denote by $\Psi_N : E \to \{1,\dots,\mathfrak{n}, N\}$ the projection defined by $\Psi_N(\eta) = x$ if $\eta \in \mathcal{E}_x$, $\Psi_N(\eta) = N$, otherwise:

$$\Psi_N(\eta) \;=\; \sum_{x\in S} x\, \mathbf{1}\{\eta \in \mathcal{E}_x\} \;+\; N\, \mathbf{1}\{\eta \in \Delta\} \,.$$

Recall from Landim (2015) the definition of the soft topology.

Theorem 2.1. *Assume that conditions* (H1)–(H3) *are in force. Fix $x \in S$ and a configuration $\eta \in \mathcal{E}_x$. Starting from η, the speeded-up, hidden Markov chain $\boldsymbol{X}_N(t) = \Psi_N\big(\eta^N(\theta_N t)\big)$ converges in the soft topology to the continuous–time Markov chain $X_{\mathcal{E}}(t)$ on $\{1,\dots,\mathfrak{n}\}$ whose jump rates are given by $r_{\mathcal{E}}(x,y)$ and which starts from x.*

This theorem is a straightforward consequence of known results. We stated it here in sake of completeness and because all the analysis of the metastable behavior of η^N_t relies on it. Its proof is presented in Section 6. We first show, using assumptions (H1) and (H2), that the trace process $\eta^{\mathcal{E}}_t$ converges in the Skorohod topology to the continuous-time Markov chain $X_{\mathcal{E}}(t)$. We combine this result with assumption (H3) to conclude the convergence of the process $\boldsymbol{X}_N(t)$ to $X_{\mathcal{E}}(t)$ in the soft topology.

One can not expect the convergence of the process $\boldsymbol{X}_N(t)$ to $X_{\mathcal{E}}(t)$ in any of the Skorohod topologies due to the very short excursions of the speeded-up process $\eta^N(t\theta_N)$ in the set Δ. To overcome this difficulty, the soft topology has been introduced.

Remark 2.2. Theorem 2.1 states that in the time scale θ_N, if we just keep track of the set $\mathcal{E}_x$ where η^N_t is and not of the specific location of the chain, we observe an evolution on the set S close to the one of a continuous-time Markov chain which jumps from x to y at rate $r_{\mathcal{E}}(x,y)$.

Remark 2.3. The function Ψ_N represents a slow variable of the chain. Indeed, we will see below that the sequence α_N^{-1} stands for the order of magnitude of the jump rates of the chain. Theorem 2.1 states that on the time scale θ_N, which is much longer than α_N, the variable $\Psi_N(\eta^N_t)$ evolves as a Markov chain. In other

words, under conditions (H1)–(H3), one still observes a Markovian dynamics after a contraction of the configuration space through the projection Ψ_N. Theorem 2.1 provides therefore a mechanism of reducing the degrees of freedom of the system, keeping the essential features of the dynamics, as the ergodic properties.

Remark 2.4. It also follows from assumptions (H1)–(H3) that the exit time from a set $\mathcal{E}_x$ is asymptotically exponential. More precisely, let $\breve{\mathcal{E}}_x$, $x \in S$, be the union of all sets $\mathcal{E}_y$ except $\mathcal{E}_x$:

$$\breve{\mathcal{E}}_x \;=\; \bigcup_{y \neq x} \mathcal{E}_y \;. \tag{2.4}$$

For every $x \in S$ and $\eta \in \mathcal{E}_x$, under $\mathbb{P}_\eta$ the distribution of $H_{\breve{\mathcal{E}}_x}/\theta_N$ converges to an exponential distribution.

Remark 2.5. Under the assumptions (H1)–(H3), the sets $\mathcal{E}_x$ are cycles in the sense of Freidlin and Wentzell (1998). More precisely, for every $x \in S$ for which $\mathcal{E}_x$ is a not a singleton, and for all $\eta \neq \xi \in \mathcal{E}_x$,

$$\lim_{N\to\infty} \mathbb{P}_\eta\big[H_\xi < H_{\breve{\mathcal{E}}_x}\big] \;=\; 1 \;.$$

This means that starting from $\eta \in \mathcal{E}_x$, the chain visits all configurations in $\mathcal{E}_x$ before hitting the set $\breve{\mathcal{E}}_x$.

2.2. *The main assumption.* We present in this subsection the main and unique hypothesis made on the sequence of Markov chains η_t^N. Fix two configurations $\eta \neq \xi \in E$. We assume that the jump rate from η to ξ is either constant equal to 0 or is always strictly positive:

$$R_N(\eta,\xi) \;=\; 0 \;\text{ for all } N \geq 1 \;\text{ or }\; R_N(\eta,\xi) \;>\; 0 \;\text{ for all } N \geq 1 \;.$$

This assumption permits to define the set of ordered bonds of E, denoted by $\mathbb{B}$, as the set of ordered pairs (η,ξ) such that $R_N(\eta,\xi) > 0$:

$$\mathbb{B} \;=\; \big\{(\eta,\xi) \in E \times E : \eta \neq \xi \,,\, R_N(\eta,\xi) > 0\big\} \;.$$

Note that the set $\mathbb{B}$ does not depend on N.

Our analysis of the metastable behavior of the sequence of Markov chain η_t^N relies on the assumption that the set of ordered bonds can be divided into equivalent classes in such a way that the all jump rates in the same equivalent class are of the same order, while the ratio between two jump rates in different classes either vanish in the limit or tend to $+\infty$. Some terminology is necessary to make this notion precise.

Ordered sequences: A set of sequences $(a_N^r : N \geq 1)$ of nonnegative real numbers, indexed by some finite set $r \in \mathfrak{R}$, is said to be *ordered* if for all $r \neq s \in \mathfrak{R}$ the sequence $\arctan\{a_N^r/a_N^s\}$ converges as $N \uparrow \infty$.

In the examples below the set $\mathfrak{R}$ will be the set of configurations E or the set of bonds $\mathbb{B}$. Let $\mathbb{Z}_+ = \{0,1,2,\dots\}$, and let $\mathfrak{A}_m$, $m \geq 1$, be the set of functions $k : \mathbb{B} \to \mathbb{Z}_+$ such that $\sum_{(\eta,\xi)\in\mathbb{B}} k(\eta,\xi) = m$.

Assumption 2.6. We assume that for every $m \geq 1$ the set of sequences

$$\Big\{ \prod_{(\eta,\xi)\in\mathbb{B}} R_N(\eta,\xi)^{k(\eta,\xi)} : N \geq 1 \Big\} \,, \quad k \in \mathfrak{A}_m$$

is ordered.

We assume from now on that the sequence of Markov chains η_t^N fulfills Assumption 2.6. In particular, the sequences $\{R_N(\eta,\xi) : N \geq 1\}$, $(\eta,\xi) \in \mathbb{B}$, are ordered.

2.3. *The shallowest valleys, the fastest slow variable.* We identify in this subsection the shortest time-scale at which a metastable behavior is observed, we introduce the shallowest valleys, and we prove that these valleys form a partition which fulfills conditions (H1)–(H3).

We first identify the valleys. Let

$$\frac{1}{\alpha_N} \;=\; \sum_{\eta\in E}\sum_{\xi:\xi\neq\eta} R_N(\eta,\xi)\;.$$

We could also have defined α_N^{-1} as $\max\{R_N(\eta,\xi) : (\eta,\xi)\in\mathbb{B}\}$. Thus, α_N^{-1} represents the order of magnitude of the largest jump rate. Denote by ξ_t^N the Markov chain η_t^N speeded-up by α_N, $\xi_t^N = \eta^N_{t\alpha_N}$, and by $R_N^\xi(\eta,\xi)$ the jump rates of the Markov chain ξ_t^N: $R_N^\xi(\eta,\xi) = \alpha_N R_N(\eta,\xi)$.

By Assumption 2.6, for every $\eta \neq \xi \in E$, $\alpha_N R_N(\eta,\xi) \to R(\eta,\xi) \in [0,1]$. In particular, the Markov chain ξ_t^N converges, in the Skorohod topology, to a Markov chain, denoted by $X_R(t)$, whose jump rates are given by $R(\eta,\xi)$. Note that this Markov chain might not be irreducible. However, by definition of α_N, there is at least one bond $(\eta,\xi)\in\mathbb{B}$ such that $R(\eta,\xi) > 0$.

Let $\lambda(\eta) = \sum_{\xi\neq\eta} R(\eta,\xi) \in \mathbb{R}_+$, and denote by E_0 the subset of points of E such that $\lambda(\eta) > 0$. For all $\eta\in E_0$, let $p(\eta,\xi) = R(\eta,\xi)/\lambda(\eta)$. It is clear that for all η, ζ in E, $\xi\in E_0$,

$$\lim_{N\to\infty} \alpha_N\,\lambda_N(\eta) \;=\lambda(\eta)\;,\quad \lim_{N\to\infty} p_N(\xi,\zeta) \;=\; p(\xi,\zeta)\;. \tag{2.5}$$

Denote by $\mathcal{E}_1,\mathcal{E}_2,\dots,\mathcal{E}_{\mathfrak{n}}$ the recurrent classes of the Markov chain $X_R(t)$, and by Δ the set of transient points, so that $\{\mathcal{E}_1,\dots,\mathcal{E}_{\mathfrak{n}},\Delta\}$ forms a partition of E:

$$E \;=\; \mathcal{E}\sqcup\Delta\;,\quad \mathcal{E} \;=\; \mathcal{E}_1\sqcup\cdots\sqcup\mathcal{E}_{\mathfrak{n}}\;. \tag{2.6}$$

Here and below we use the notation $\mathcal{A}\sqcup\mathcal{B}$ to represent the union of two disjoint sets $\mathcal{A}$, $\mathcal{B}$: $\mathcal{A}\sqcup\mathcal{B} = \mathcal{A}\cup\mathcal{B}$, and $\mathcal{A}\cap\mathcal{B} = \varnothing$.

Note that the sets $\mathcal{E}_x$, $x\in S = \{1,\dots,\mathfrak{n}\}$, do not depend on N. If $\mathfrak{n} = 1$, the chain does not possess valleys. This is the case, for instance, if the rates $R_N(x,y)$ are independent of N. Assume, therefore, and up to the end of this subsection, that $\mathfrak{n}\geq 2$.

Since the set $\mathcal{E}_x$ is a recurrent class for the chain $X_R(t)$, and since the process ξ_t^N converges to $X_R(t)$, the chain ξ_t^N does not leave the set $\mathcal{E}_x$ if it starts there. This means that the chain η_t^N remains in the set $\mathcal{E}_x$ in the time-scale α_N if it starts from a point in this set. That is, the set $\mathcal{E}_x$ is a well for the chain η_t^N.

To observe the chain leaving the set $\mathcal{E}_x$, one has to wait a much longer time. The correct time-scale for this event is given by $\mu_N(\mathcal{E}_x)/\mathrm{cap}_N(\mathcal{E}_x,\breve{\mathcal{E}}_x)$. In particular, a metastable behavior of the chain η_t^N will be observed in the smallest of these time-scales. For this reason, let θ_N be defined by

$$\frac{1}{\theta_N} \;=\; \sum_{x\in S} \frac{\mathrm{cap}_N(\mathcal{E}_x,\breve{\mathcal{E}}_x)}{\mu_N(\mathcal{E}_x)}\;. \tag{2.7}$$

Since the sum is of the order of the maximum, θ_N is of the order of the smallest of the sequences $\mu_N(\mathcal{E}_x)/\mathrm{cap}_N(\mathcal{E}_x, \check{\mathcal{E}}_x)$, and, as explained above, it is on this time-scale that a metastable behavior is expected to occur.

Theorem 2.7. *The partition $\mathcal{E}_1, \dots, \mathcal{E}_{\mathfrak{n}}, \Delta$ and the time scales α_N, θ_N fulfill the conditions* (H1)–(H3). *Moreover, For every $x \in S$ and every $\eta \in \mathcal{E}_x$, there exists $m_x(\eta) \in (0,1]$ such that*

$$\lim_{N\to\infty} \frac{\mu_N(\eta)}{\mu_N(\mathcal{E}_x)} \;=\; m_x(\eta)\,. \tag{H0}$$

Remark 2.8. The jump rates $r_{\mathcal{E}}(x,y)$ which appear in condition (H1) are introduced in Lemma 7.1. It follows from Theorems 2.1 and 2.7 that in the time-scale θ_N the chain η_t^N evolves among the sets $\mathcal{E}_x$, $x \in S$, as a Markov chain which jumps from x to y at rate $r_{\mathcal{E}}(x,y)$.

In the next three remarks we present some outcomes of Theorem 2.1 and 2.7 on the evolution of the chain η_t^N in a time-scale longer than θ_N. These remarks anticipate the recursive procedure of the next subsection.

Remark 2.9. The jump rates $r_{\mathcal{E}}(x,y)$ define a Markov chain on S, represented by $X_{\mathcal{E}}(t)$. Denote by T the set of transient points of this chain and assume that $T \neq \varnothing$. It follows from Theorem 2.1 that in the time-scale θ_N, starting from a set $\mathcal{E}_x$, $x \in T$, the chain η_t^N leaves the set $\mathcal{E}_x$ at an asymptotically exponential time, and never returns to $\mathcal{E}_x$ after a finite number of visits to this set. In particular, if we observe the chain η_t^N in a longer time-scale than θ_N, starting from $\mathcal{E}_x$ the chain remains only a negligible amount of time at $\mathcal{E}_x$.

Remark 2.10. Denote by A the set of absorbing points of $X_{\mathcal{E}}(t)$, and assume that $A \neq \varnothing$. In this case, in the time-scale θ_N, starting from a set $\mathcal{E}_x$, $x \in A$, the chain η_t^N never leaves the set $\mathcal{E}_x$. To observe a non-trivial behavior starting from this set one has to consider longer-time scales.

Remark 2.11. Finally, denote by $\mathcal{C}_1, \dots, \mathcal{C}_{\mathfrak{p}}$ the equivalent classes of $X_{\mathcal{E}}(t)$. Suppose that there is a class, say $\mathcal{C}_1$, of recurrent points which is not a singleton. In this case, starting from a set $\mathcal{E}_x$, $x \in \mathcal{C}_1$, in the time-scale θ_N, the chain η_t^N leaves the set $\mathcal{E}_x$ at an asymptotically exponential time, and returns to $\mathcal{E}_x$ infinitely many times.

Suppose now that there are at least two classes, say $\mathcal{C}_1$ and $\mathcal{C}_2$, of recurrent points. This means that in the time-scale θ_N, starting from a set $\mathcal{E}_x$, $x \in \mathcal{C}_1$, the process never visits a set $\mathcal{E}_y$ for $y \in \mathcal{C}_2$. For this to occur one has to observe the chain η_t^N in a longer time-scale.

Denote by $R_1, \dots, R_{\mathfrak{m}}$ the recurrent classes of $X_{\mathcal{E}}(t)$. In the next subsection, we derive a new time-scale at which one observes jumps from sets of the form $\mathcal{F}_a = \cup_{x\in R_a} \mathcal{E}_x$ to sets of the form $\mathcal{F}_b = \cup_{x\in R_b} \mathcal{E}_x$.

2.4. *The deep valleys and slow variables.* We obtained in the previous subsection two time-scales α_N, θ_N, and a partition $\mathcal{E}_1, \dots, \mathcal{E}_{\mathfrak{n}}, \Delta$ of the state space E which satisfy conditions (H0)–(H3). We present in this subsection a recursive procedure. Starting from two time-scales β_N^-, β_N, and a partition $\mathcal{F}_1, \dots, \mathcal{F}_{\mathfrak{p}}, \Delta_{\mathcal{F}}$ of the state space E satisfying the assumptions (H0)–(H3) and such that $\mathfrak{p} \geq 2$, it provides a longer time-scale β_N^+ and a coarser partition $\mathcal{G}_1, \dots, \mathcal{G}_{\mathfrak{q}}, \Delta_{\mathcal{G}}$ which fulfills conditions (H0)–(H3) with respect to the sequences β_N, β_N^+.

Consider a partition $\mathcal{F}_1, \dots, \mathcal{F}_{\mathfrak{p}}, \Delta_{\mathcal{F}}$ of the set E and two sequences β_N^-, β_N such that $\beta_N^-/\beta_N \to 0$. Assume that $\mathfrak{p} \geq 2$ and that the partition and the sequences β_N^-, β_N satisfy conditions (H0)–(H3). Denote by $r_{\mathcal{F}}(x,y)$ the jump rates appearing in assumption (H1).

The coarser partition. Let $P = \{1, \dots, \mathfrak{p}\}$ and let $X_{\mathcal{F}}(t)$ be the P-valued Markov chain whose jumps rates are given by $r_{\mathcal{F}}(x,y)$.

Denote by $G_1, G_2, \dots, G_{\mathfrak{q}}$ the recurrent classes of the chain $X_{\mathcal{F}}(t)$, and by $G_{\mathfrak{q}+1}$ the set of transient points. The sets $G_1, \dots, G_{\mathfrak{q}+1}$ form a partition of P. We claim that $\mathfrak{q} < \mathfrak{p}$. Fix $x \in P$ such that $\sum_{y \neq x} r_{\mathcal{F}}(x,y) > 0$, whose existence is guaranteed by hypothesis (H1). Suppose that the point x is transient. In this case the number of recurrent classes must be smaller than $\mathfrak{p}$. If, on the other hand, x is recurrent, the recurrent class which contains x must have at least two elements, and the number of recurrent classes must be smaller than $\mathfrak{p}$.

Let $Q = \{1, \dots, \mathfrak{q}\}$,

$$\mathcal{G}_a \;=\; \bigcup_{x \in G_a} \mathcal{F}_x \;, \quad \Delta_* \;=\; \bigcup_{x \in G_{\mathfrak{q}+1}} \mathcal{F}_x \;, \quad \Delta_{\mathcal{G}} = \Delta_{\mathcal{F}} \cup \Delta_* \;, \quad a \in Q \;. \tag{2.8}$$

Since, by (2.6), $\{\mathcal{F}_1, \dots, \mathcal{F}_{\mathfrak{p}}, \Delta_{\mathcal{F}}\}$ forms a partition of E, $\{\mathcal{G}_1, \dots, \mathcal{G}_{\mathfrak{q}}, \Delta_{\mathcal{G}}\}$ also forms a partition of E:

$$E \;=\; \mathcal{G} \sqcup \Delta_{\mathcal{G}} \;, \quad \mathcal{G} \;=\; \mathcal{G}_1 \sqcup \cdots \sqcup \mathcal{G}_{\mathfrak{q}} \;. \tag{2.9}$$

The longer time-scale. For $a \in Q = \{1, \dots, \mathfrak{q}\}$, let $\breve{\mathcal{G}}_a$ be the union of all sets $\mathcal{G}_b$ except $\mathcal{G}_a$:

$$\breve{\mathcal{G}}_a \;=\; \bigcup_{b \neq a} \mathcal{G}_b \;.$$

Assume that $\mathfrak{q} > 1$, and let β_N^+ be given by

$$\frac{1}{\beta_N^+} \;=\; \sum_{a \in Q} \frac{\mathrm{cap}_N(\mathcal{G}_a, \breve{\mathcal{G}}_a)}{\mu_N(\mathcal{G}_a)} \;. \tag{2.10}$$

Theorem 2.12. *The partition $\mathcal{G}_1, \dots, \mathcal{G}_q$, $\Delta_{\mathcal{G}}$ and the time scales (β_N, β_N^+) satisfy conditions* (H0)–(H3).

Remark 2.13. It follows from Theorems 2.1 and 2.12 that the chain η_t^N exhibits a metastable behavior in the time-scale β_N^+ if $\mathfrak{q} > 1$. We refer to Remarks 2.2, 2.3, 2.4 and 2.5.

Remark 2.14. As $\mathfrak{q} < \mathfrak{p}$ and as we need $\mathfrak{p}$ to be greater than or equal to 2 to apply the iterative procedure, this recursive algorithm ends after a finite number of steps.

If $\mathfrak{q} = 1$, β_N is the longest time-scale at which a metastable behavior is observed. In this time-scale, the chain η_t^N jumps among the sets $\mathcal{F}_x$ as does the chain $X_{\mathcal{F}}(t)$ until it reaches the set $\mathcal{G}_1 = \cup_{x \in G_1} \mathcal{F}_x$. Once in this set, it remains there for ever jumping among the sets $\mathcal{F}_x$, $x \in G_1$, as the Markov chain $X_{\mathcal{F}}(t)$, which restricted to G_1 is an irreducible Markov chain.

The successive valleys: Observe that the valleys $\mathcal{G}_a$ were obtained as the recurrent classes of the Markov chain $X_{\mathcal{F}}(t)$: $\mathcal{G}_a = \cup_{x \in G_a} \mathcal{F}_x$, where G_a is a recurrent class of $X_{\mathcal{F}}(t)$. In particular, at any time-scale the valleys are formed by unions of the valleys obtained in the first step of the recursive argument, which were denoted by $\mathcal{E}_x$ in the previous subsection. Moreover, by (H0), each configuration in $\mathcal{G}_a$ has measure of the same order.

Remark 2.15. The recursive procedure (2.8) creates a rooted tree, constructed from the leaves to the root. All vertices of this tree are subsets of E, each generation forms a partition of the set E, and the sets $\mathcal{E}_1, \dots, \mathcal{E}_{\mathfrak{n}}, \Delta$, introduced in Subsection 2.3, compose the leaves of the tree. Equation (2.8) describes the parents of a generation: If $\mathcal{F}_1, \dots, \mathcal{F}_{\mathfrak{p}}, \Delta_{\mathcal{F}}$ represent the vertices of the k-th generation, the sets $\mathcal{G}_1, \dots, \mathcal{G}_{\mathfrak{q}}, \Delta_{\mathcal{G}}$ given by (2.8) form the generation $k-1$. The iteration ends, with two sets $\mathcal{G}_1$ and $\Delta_{\mathcal{G}}$, when the set Q, introduced just before (2.8), is a singleton. To complete the construction of the tree we define the set E as the root.

Observe that the length of the rooted tree corresponds to the number of time-scales at which a metastable behavior is observed minus 1, and that each generation is formed by disjoint subsets of the previous generation. In particular, the generations form a sequence of partitions of E, strictly finer at each step.

Remark 2.16. Scoppola (1993) and Olivieri and Scoppola (1995, 1996) proposed a similar construction for discrete-time Markov chains assuming that the jump rates $p_N(\eta, \xi)$, η, $\xi \in E$, of the chain satisfy

$$\lim_{N\to\infty} \frac{1}{N} \log p_N(\eta, \xi) \;=\; \Delta(\eta, \xi)$$

for some non-negative function $\Delta : \mathbb{B} \to \mathbb{R}_+$. This condition is replaced here by Assumption 2.6.

Conclusion: We presented an iterative method which provides a finite sequence of time-scales and of partitions of the set E satisfying conditions (H0)–(H3). At each step, the time scales become longer and the partitions coarser. By Theorem 2.1, to each pair of time-scale and partition corresponds a metastable behavior of the chain η_t^N. This recursive algorithm provides all time-scales at which a metastable behavior of the chain η_t^N is observed, and all slow variables associated to the dynamics.

3. What do we learn from Assumption 2.6?

We prove in this section that the jump rates of the trace processes satisfy Assumption 2.6, and that some sequences, such as the one formed by the measures of the configurations, are ordered.

Assertion 3.A. Let F be a proper subset of E and denote by $R_N^F(\eta, \xi)$, $\eta \neq \xi \in F$, the jump rates of the trace of η_t^N on F. The jump rates $R_N^F(\eta, \xi)$ satisfy Assumption 2.6.

Proof: We prove this assertion by removing one by one the elements of $E \setminus F$. Assume that $F = E \setminus \{\zeta\}$ for some $\zeta \in E$. By Corollary 6.2 in Beltrán and Landim (2010) and by the equation following the proof of this corollary, for $\eta \neq \xi \in F$, $R_N^F(\eta, \xi) = R_N(\eta, \xi) + R_N(\eta, \zeta) p_N(\zeta, \xi)$. Hence,

$$R_N^F(\eta, \xi) \;=\; \frac{\sum_{w\in E} R_N(\eta, \xi) R_N(\zeta, w) + R_N(\eta, \zeta) R_N(\zeta, \xi)}{\sum_{w\in E} R_N(\zeta, w)} \,. \tag{3.1}$$

It is easy to check from this identity that Assumption 2.6 holds for the jump rates R_N^F. It remains to proceed recursively to complete the proof. □

Lemma 3.1. *The sequences* $\{\mu_N(\eta) : N \geq 1\}$, $\eta \in E$, *are ordered.*

Proof: Fix $\eta \neq \xi \in E$ and let $F = \{\eta, \xi\}$. By Beltrán and Landim (2010, Proposition 6.3), the stationary state of the trace of η_t^N on F, denoted by μ_N^F, is given by $\mu_N^F(\eta) = \mu_N(\eta)/\mu_N(F)$. As μ_N^F is the invariant probability measure, $\mu_N^F(\eta)R_N^F(\eta,\xi) = \mu_N^F(\xi)R_N^F(\xi,\eta)$. Therefore, $\mu_N(\eta)/\mu_N(\xi) = \mu_N^F(\eta)/\mu_N^F(\xi) = R_N^F(\xi,\eta)/R_N^F(\eta,\xi)$. By Assertion 3.A, the sequences $\{R_N^F(a,b) : N \geq 1\}$, $a \neq b \in \{\eta, \xi\}$ are ordered. This completes the proof of the lemma. □

The previous lemma permits to divide the configurations of E into equivalent classes by declaring η equivalent to η', $\eta \sim \eta'$, if $\mu_N(\eta)/\mu_N(\eta')$ converges to a real number belonging to $(0, \infty)$.

Assertion 3.B. Let F be a proper subset of E. For every bond $(\eta', \xi') \in \mathbb{B}$ and every $m \geq 1$ the set of sequences

$$\Big\{ \prod_{(\eta,\xi)\in\mathbb{B}} R_N^F(\eta,\xi)^{k(\eta,\xi)} R_N(\eta',\xi') : N \geq 1 \Big\}, \quad k \in \mathfrak{A}_m$$

is ordered.

Proof: We proceed as in the proof of Assertion 3.A, by removing one by one the elements of $E \setminus F$. Fix $\zeta \in E \setminus F$. It follows from (3.1) and from Assumption 2.6 that the claim of the assertion holds for $F' = E \setminus \{\zeta\}$.

Fix $\zeta' \in E \setminus F$, $\zeta' \neq \zeta$. By using formula (3.1), to express the rates $R^{E\setminus\{\zeta,\zeta'\}}$ in terms of the rates $R^{E\setminus\{\zeta\}}$, and the statement of this assertion for $F' = E \setminus \{\zeta\}$ we prove that this assertion also holds for $F' = E \setminus \{\zeta, \zeta'\}$. Iterating this algorithm we complete the proof of the assertion. □

Denote by $c_N(\eta,\xi) = \mu_N(\eta)R_N(\eta,\xi)$, $(\eta,\xi) \in \mathbb{B}$, the (generally asymmetric) conductances.

Lemma 3.2. *The conductances $\{c_N(\eta,\xi) : N \geq 1\}$, $(\eta,\xi) \in \mathbb{B}$, are ordered.*

Proof: Consider two bonds (η,ξ), (η',ξ') in $\mathbb{B}$. As in the proof of Lemma 3.1, we may express the ratio of the conductances as

$$\frac{c_N(\eta,\xi)}{c_N(\eta',\xi')} = \frac{\mu_N(\eta)R_N(\eta,\xi)}{\mu_N(\eta')R_N(\eta',\xi')} = \frac{R_N^F(\eta',\eta)R_N(\eta,\xi)}{R_N^F(\eta,\eta')R_N(\eta',\xi')},$$

where $F = \{\eta, \eta'\}$. It remains to recall the statement of assertion 3.B to complete the proof of the lemma. □

Denote by $\mathbb{B}^s$ the symmetrization of the set $\mathbb{B}$, that is, the set of bonds (η,ξ) such that (η,ξ) or (ξ,η) belongs to $\mathbb{B}$:

$$\mathbb{B}^s = \big\{(\eta,\xi) \in E \times E : \eta \neq \xi\,,\ (\eta,\xi) \in \mathbb{B} \ \text{ or } \ (\xi,\eta) \in \mathbb{B}\big\}\,.$$

Denote by $c_N^s(\eta,\xi)$, $(\eta,\xi) \in \mathbb{B}^s$, the symmetric part of the conductance:

$$c_N^s(\eta,\xi) = \frac{1}{2}\big\{c_N(\eta,\xi) + c_N(\xi,\eta)\big\}\,. \tag{3.2}$$

Next result is a straightforward consequence of the previous lemma.

Corollary 3.3. *The symmetric conductances $\{c_N^s(\eta,\xi) : N \geq 1\}$, $(\eta,\xi) \in \mathbb{B}^s$, are ordered.*

As in Lemma 3.1, the previous corollary permits to divide the set $\mathbb{B}^s$ into equivalent classes by declaring (η,ξ) equivalent to (η',ξ'), $(\eta,\xi) \sim (\eta',\xi')$, if $c_N^s(\eta,\xi)/c_N^s(\eta',\xi')$ converges to a number in $(0,\infty)$.

It is possible to deduce from Assumption 2.6 that many other sequences are ordered. We do not present these results here as we do not use them below.

4. Cycles, sector condition and capacities

We prove in this section that the generator of a Markov chain on a finite set can be decomposed as the sum of cycle generators and that it satisfies a sector condition. This last bound permits to estimate the capacity between two sets by the capacity between the same sets for the reversible process.

Throughout this section, E is a fixed finite set and $\mathcal{L}$ represents the generator of an E-valued, continuous-time Markov chain. We adopt all notation introduced in Section 2, removing the index N since the chain is fixed. We start with some definitions.

In a finite set, the decomposition of a generator into cycle generators is very simple. The problem for infinite sets is much more delicate. We refer to Gabrielli and Valente (2012) for a discussion of the question.

Cycle: A cycle is a sequence of distinct configurations $(\eta_0,\eta_1,\dots,\eta_{n-1},\eta_n=\eta_0)$ whose initial and final configuration coincide: $\eta_i \neq \eta_j \in E$, $i \neq j \in \{0,\dots,n-1\}$. The number n is called the length of the cycle.

Cycle generator: A generator $\mathcal{L}$ of an E-valued Markov chain, whose jump rates are denoted by $R(\eta,\xi)$, is said to be a cycle generator associated to the cycle $\mathfrak{c}=(\eta_0,\eta_1,\dots,\eta_{n-1},\eta_n=\eta_0)$ if there exists reals $r_i>0$, $0\le i<n$, such that

$$R(\eta,\xi) \;=\; \begin{cases} r_i & \text{if } \eta=\eta_i \text{ and } \xi=\eta_{i+1} \text{ for some } 0\le i<n\,,\\ 0 & \text{otherwise}\,. \end{cases}$$

We denote this cycle generator by $\mathcal{L}_{\mathfrak{c}}$. Note that

$$(\mathcal{L}_{\mathfrak{c}} f)(\eta) \;=\; \sum_{i=0}^{n-1} \mathbf{1}\{\eta=\eta_i\}\, r_i\, [f(\eta_{i+1})-f(\eta_i)]\,.$$

Sector condition: A generator $\mathcal{L}$ of an E-valued, irreducible Markov chain, whose unique invariant probability measure is denoted by μ, is said to satisfy a sector condition if there exists a constant $C_0<\infty$ such that for all functions $f,g:E\to\mathbb{R}$,

$$\langle \mathcal{L}f, g\rangle_\mu^2 \;\le\; C_0 \langle (-\mathcal{L}f), f\rangle_\mu \, \langle (-\mathcal{L}g), g\rangle_\mu\,.$$

In this formula, $\langle f,g\rangle_\mu$ represents the scalar product in $L^2(\mu)$:

$$\langle f,g\rangle_\mu \;=\; \sum_{\eta\in E} f(\eta)\, g(\eta)\, \mu(\eta)\,.$$

We claim that every cycle generator satisfies a sector condition and that every generator $\mathcal{L}$ of an E-valued Markov chain, stationary with respect to a probability measure μ, can be decomposed as the sum of cycle generators which are stationary with respect to μ.

Assertion 4.A. Consider a cycle $\mathfrak{c} = (\eta_0, \eta_1, \dots, \eta_{n-1}, \eta_n = \eta_0)$ of length $n \geq 2$ and let $\mathcal{L}$ be a cycle generator associated to $\mathfrak{c}$. Denote the jump rates of $\mathcal{L}$ by $R(\eta_i, \eta_{i+1})$. A measure μ is stationary for $\mathcal{L}$ if and only if

$$\mu(\eta_i)\, R(\eta_i, \eta_{i+1}) \ \text{ is constant .} \tag{4.1}$$

The proof of the previous assertion is elementary and left to the reader. The proof of the next one can be found in Komorowski et al. (2012, Lemma 5.5.8).

Assertion 4.B. Let $\mathcal{L}$ be a cycle generator associated to a cycle $\mathfrak{c}$ of length n. Then, $\mathcal{L}$ satisfies a sector condition with constant $2n$: For all f, $g : E \to \mathbb{R}$,

$$\langle \mathcal{L} f, g\rangle_\mu^2 \ \leq \ 2n \, \langle (-\mathcal{L} f), f\rangle_\mu \, \langle (-\mathcal{L} g), g\rangle_\mu \ .$$

Lemma 4.1. *Let $\mathcal{L}$ be a generator of an E-valued, irreducible Markov chain. Denote by μ the unique invariant probability measure. Then, there exists cycles $\mathfrak{c}_1, \dots, \mathfrak{c}_p$ such that*

$$\mathcal{L} \ = \ \sum_{j=1}^{p} \mathcal{L}_{\mathfrak{c}_j} \ ,$$

where $\mathcal{L}_{\mathfrak{c}_j}$ are cycle generators associated to $\mathfrak{c}_j$ which are stationary with respect to μ.

Proof: The proof consists in eliminating successively all 2-cycles (cycles of length 2), then all 3-cycles and so on up to the $|E|$-cycle if there is one left. Denote by $R(\eta, \xi)$ the jump rates of the generator $\mathcal{L}$ and by $\mathbb{C}_2$ the set of all 2-cycles (η, ξ, η) such that $R(\eta, \xi) R(\xi, \eta) > 0$. Note that the cycle (η, ξ, η) coincide with the cycle (ξ, η, ξ).

Fix a cycle $\mathfrak{c} = (\eta, \xi, \eta) \in \mathbb{C}_2$. Let $\bar{c}(\eta, \xi) = \min\{\mu(\eta) R(\eta, \xi), \mu(\xi) R(\xi, \eta)\}$ be the minimal conductance of the edge (η, ξ), and let $R_{\mathfrak{c}}(\eta, \xi)$ be the jump rates given by $R_{\mathfrak{c}}(\eta, \xi) = \bar{c}(\eta, \xi)/\mu(\eta)$, $R_{\mathfrak{c}}(\xi, \eta) = \bar{c}(\eta, \xi)/\mu(\xi)$. Observe that $R_{\mathfrak{c}}(\zeta, \zeta') \leq R(\zeta, \zeta')$ for all (ζ, ζ'), and that $R_{\mathfrak{c}}(\xi, \eta) = R(\xi, \eta)$ or $R_{\mathfrak{c}}(\eta, \xi) = R(\eta, \xi)$.

Denote by $\mathcal{L}_{\mathfrak{c}}$ the generator associated the the jump rates $R_{\mathfrak{c}}$. Since $\mu(\eta) R_{\mathfrak{c}}(\eta, \xi) = \bar{c}(\eta, \xi) = \mu(\xi) R_{\mathfrak{c}}(\xi, \eta)$, by (4.1), μ is a stationary state for $\mathcal{L}_{\mathfrak{c}}$ (actually, reversible). Let $\mathcal{L}_1 = \mathcal{L} - \mathcal{L}_{\mathfrak{c}}$ so that

$$\mathcal{L} \ = \ \mathcal{L}_1 \ + \ \mathcal{L}_{\mathfrak{c}} \ .$$

As $R_{\mathfrak{c}}(\zeta, \zeta') \leq R(\zeta, \zeta')$, $\mathcal{L}_1$ is the generator of a Markov chain. Since both $\mathcal{L}$ and $\mathcal{L}_{\mathfrak{c}}$ are stationary for μ, so is $\mathcal{L}_1$. Finally, if we draw an arrow from ζ to ζ' if the jump rate from ζ to ζ' is strictly positive, the number of arrows for the generator $\mathcal{L}_1$ is equal to the number of arrows for the generator $\mathcal{L}$ minus 1 or 2. This procedure has therefore strictly decreased the number of arrows of $\mathcal{L}$.

We may repeat the previous algorithm to $\mathcal{L}_1$ to remove from $\mathcal{L}$ all 2-cycles (η, ξ, η) such that $R(\eta, \xi) R(\xi, \eta) > 0$. Once this has been accomplished, we may remove all 3-cycles $(\eta_0, \eta_1, \eta_2, \eta_3 = \eta_0)$ such that $\prod_{0 \leq i < 3} R(\eta_i, \eta_{i+1}) > 0$. At each step at least one arrow is removed from the generator which implies that after a finite number of steps all 3-cycles are removed.

Once all k-cycles have been removed, $2 \leq k < |E|$, we have obtained a decomposition of $\mathcal{L}$ as

$$\mathcal{L} \ = \ \sum_{k=2}^{|E|-1} \mathcal{L}_k \ + \ \hat{\mathcal{L}} \ ,$$

where $\mathcal{L}_k$ is the sum of k-cycle generators and is stationary with respect to μ, and $\hat{\mathcal{L}}$ is a generator, stationary with respect to μ, and with no k-cycles, $2 \le k < |E|$. If $\hat{\mathcal{L}}$ has an arrow, as it is stationary with respect to μ and has no k-cycles, $\hat{\mathcal{L}}$ must be an $|E|$-cycle generator, providing the decomposition stated in the lemma. □

Remark 4.2. Observe that a generator $\mathcal{L}$ is reversible with respect to μ if and only if it has a decomposition in 2-cycles. Given a measure μ on a finite state space, for example the Gibbs measure associated to a Hamiltonian at a fixed temperature, by introducing k-cycles satisfying (4.1) it is possible to define non-reversible dynamics which are stationary with respect to μ. The previous lemma asserts that this is the only way to define such dynamics.

Corollary 4.3. *The generator $\mathcal{L}$ satisfies a sector condition with constant bounded by $2|E|$: For all f, $g : E \to \mathbb{R}$,*

$$\langle \mathcal{L} f, g\rangle_\mu^2 \;\le\; 2|E|\, \langle (-\mathcal{L} f), f\rangle_\mu \, \langle (-\mathcal{L} g), g\rangle_\mu \,.$$

Proof: Fix f and $g : E \to \mathbb{R}$. By Lemma 4.1,

$$\langle \mathcal{L} f, g\rangle_\mu^2 \;=\; \Big(\sum_{j=1}^p \langle \mathcal{L}_{\mathfrak{c}_j} f, g\rangle_\mu \Big)^2 \,,$$

where $\mathcal{L}_{\mathfrak{c}_j}$ is a cycle generator, stationary with respect to μ, associated to the cycle $\mathfrak{c}_j$. By Assertion 4.B and by Schwarz inequality, since all cycles have length at most $|E|$, the previous sum is bounded by

$$2|E| \sum_{j=1}^p \langle (-\mathcal{L}_{\mathfrak{c}_j} f), f\rangle_\mu \sum_{k=1}^p \langle (-\mathcal{L}_{\mathfrak{c}_k} g), g\rangle_\mu \;=\; 2|E|\, \langle (-\mathcal{L} f), f\rangle_\mu \, \langle (-\mathcal{L} g), g\rangle_\mu \,,$$

as claimed □

Denote by $R^s(\eta, \xi)$ the symmetric part of the jump rates $R(\eta, \xi)$:

$$R^s(\eta, \xi) \;=\; \frac{1}{2} \Big\{ R(\eta, \xi) \;+\; \frac{\mu(\xi)}{\mu(\eta)}\, R(\xi, \eta) \Big\} \,. \tag{4.2}$$

Denote by η_t^s the E-valued Markov chain whose jump rates are given by R^s. The chain η_t^s is called the reversible chain.

For two disjoint subsets A, B of E, denote by $\mathrm{cap}(A, B)$ (resp. $\mathrm{cap}^s(A, B)$) the capacity between A and B (for the reversible chain). Next result follows from Corollary 4.3 and Lemmas 2.5 and 2.6 in Gaudillière and Landim (2014).

Corollary 4.4. *Fix two disjoint subsets A, B of E. Then,*

$$\mathrm{cap}^s(A, B) \;\le\; \mathrm{cap}(A, B) \;\le\; 2|E|\, \mathrm{cap}^s(A, B) \,.$$

We conclude the section with an identity and an inequality which will be used several times in this article. Let A and B be two disjoint subsets of E. By definition of the capacity

$$\mathrm{cap}(A, B) \;=\; \sum_{\eta \in A} \mu(\eta)\, \lambda(\eta)\, \mathbb{P}_\eta \big[H_B < H_A^+ \big] \;=\; \sum_{\eta \in A} \mu(\eta)\, \lambda(\eta) \sum_{\xi \in B} \mathbb{P}_\eta \big[H_\xi = H_{A \cup B}^+ \big] \,.$$

Therefore, if we denote by $R^{A \cup B}(\eta, \xi)$, $\eta \neq \xi \in A \cup B$, the jump rates of the trace of the chain η_t on the set $A \cup B$, by Beltrán and Landim (2010, Proposition 6.1),

$$\mathrm{cap}(A, B) \;=\; \sum_{\eta \in A} \mu(\eta) \sum_{\xi \in B} R^{A \cup B}(\eta, \xi) \,. \tag{4.3}$$

Let A be a non-empty subset of E and denote by $R^A(\eta,\xi)$ the jump rates of the trace of η_t on A. We claim that for all $\eta \neq \xi \in A$,

$$\mu(\eta)\, R^A(\eta,\xi) \;\le\; \mathrm{cap}(\eta,\xi)\,. \tag{4.4}$$

Denote by $\lambda^A(\zeta)$ the holding rates of the trace process on A and by $p^A(\zeta,\zeta')$ the jump probabilities. By definition,

$$R^A(\eta,\xi) \;=\; \lambda^A(\eta)\, p^A(\eta,\xi) \;=\; \lambda^A(\eta)\, \mathbb{P}_\eta[H_\xi = H_A^+] \;\le\; \lambda^A(\eta)\, \mathbb{P}_\eta[H_\xi < H_\eta^+]\,.$$

Multiplying both sides of this inequality by $\mu_A(\eta) = \mu(\eta)/\mu(A)$, by definition of the capacity we obtain that

$$\mu_A(\eta)\, R^A(\eta,\xi) \;\le\; \mathrm{cap}_A(\eta,\xi)\,,$$

where $\mathrm{cap}_A(\eta,\xi)$ stands for the capacity with respect to the trace process on A. To complete the proof of (4.4), it remains to recall formula (A.10) in Beltrán and Landim (2012).

5. Reversible chains and capacities

We present in this section some estimates for the capacity of reversible, finite state Markov chains obtained in Beltrán and Landim (2011). These results are useful since we proved in Corollary 4.4 that the capacity between two disjoint subsets $\mathcal{A}$, $\mathcal{B}$ of E is of the same order as the capacity with respect to the reversible chain.

Recall from (3.2) that we denote by $c_N^s(\eta,\xi)$ the symmetric conductance of the bond (η,ξ). Fix two disjoint subsets $\mathcal{A}$, $\mathcal{B}$ of E. A self-avoiding path γ from $\mathcal{A}$ to $\mathcal{B}$ is a sequence of configurations $(\eta_0,\eta_1,\dots,\eta_n)$ such that $\eta_0 \in \mathcal{A}$, $\eta_n \in \mathcal{B}$, $\eta_i \neq \eta_j$, $i \neq j$, $c_N^s(\eta_i,\eta_{i+1}) > 0$, $0 \le i < n$. Denote by $\Gamma_{\mathcal{A},\mathcal{B}}$ the set of self-avoiding paths from $\mathcal{A}$ to $\mathcal{B}$ and let

$$\boldsymbol{c}_N^s(\gamma) \;=\; \min_{0\le i<n} c_N^s(\eta_i,\eta_{i+1})\,, \quad \boldsymbol{c}_N^s(\mathcal{A},\mathcal{B}) \;=\; \max_{\gamma\in\Gamma_{\mathcal{A},\mathcal{B}}} \boldsymbol{c}_N^s(\gamma)\,. \tag{5.1}$$

For two configurations η, ξ, we represent $\boldsymbol{c}_N^s(\{\eta\},\{\xi\})$ by $\boldsymbol{c}_N^s(\eta,\xi)$. Note that $\boldsymbol{c}_N^s(\eta,\xi) \le c_N^s(\eta,\xi)$, with possibly a strict inequality.

Fix two disjoint subsets $\mathcal{A}$, $\mathcal{B}$ of E and a configuration $\eta \notin \mathcal{A}\cup\mathcal{B}$. We claim that

$$\boldsymbol{c}_N^s(\mathcal{A},\mathcal{B}) \;\ge\; \min\{\boldsymbol{c}_N^s(\mathcal{A},\eta)\,,\, \boldsymbol{c}_N^s(\eta,\mathcal{B})\}\,. \tag{5.2}$$

Indeed, there exist a self-avoiding path γ_1 from $\mathcal{A}$ to η, and a self-avoiding path γ_2 from η to $\mathcal{B}$ such that $\boldsymbol{c}_N^s(\mathcal{A},\eta) = \boldsymbol{c}_N^s(\gamma_1)$, $\boldsymbol{c}_N^s(\eta,\mathcal{B}) = \boldsymbol{c}_N^s(\gamma_2)$. Juxtaposing the paths γ_1 and γ_2, we obtain a path γ from $\mathcal{A}$ to $\mathcal{B}$. Of course, the path γ may not be self-avoiding, may return to $\mathcal{A}$ before reaching $\mathcal{B}$, or may reach $\mathcal{B}$ before hitting η. In any case, we may obtain from γ a subpath $\hat\gamma$ which is self-avoiding and which connects $\mathcal{A}$ to $\mathcal{B}$. Subpath in the sense that all bonds (η_i,η_{i+1}) which appear in $\hat\gamma$ also appear in γ. In particular,

$$\boldsymbol{c}_N^s(\hat\gamma) \;\ge\; \boldsymbol{c}_N^s(\gamma) \;=\; \min\{\boldsymbol{c}_N^s(\gamma_1)\,,\, \boldsymbol{c}_N^s(\gamma_2)\} \;=\; \min\{\boldsymbol{c}_N^s(\mathcal{A},\eta)\,,\, \boldsymbol{c}_N^s(\eta,\mathcal{B})\}\,.$$

To complete the proof of claim (5.2), it remains to observe that $\boldsymbol{c}_N^s(\mathcal{A},\mathcal{B}) \ge \boldsymbol{c}_N^s(\hat\gamma)$.

Fix two disjoint subsets $\mathcal{A}$, $\mathcal{B}$ of E and configurations $\eta_i \notin \mathcal{A}\cup\mathcal{B}$, $1 \le i \le n$, such that $\eta_i \neq \eta_j$, $i \neq j$. Iterating inequality (5.2) we obtain that

$$\boldsymbol{c}_N^s(\mathcal{A},\mathcal{B}) \;\ge\; \min\{\boldsymbol{c}_N^s(\mathcal{A},\eta_1)\,,\, \boldsymbol{c}_N^s(\eta_1,\eta_2)\,,\,\dots\,,\, \boldsymbol{c}_N^s(\eta_{n-1},\eta_n)\,,\, \boldsymbol{c}_N^s(\eta_n,\mathcal{B})\}\,. \tag{5.3}$$

We conclude this section relating the symmetric capacity between two sets $\mathcal{A}$, $\mathcal{B}$ of E to the symmetric conductances $\boldsymbol{c}_N^s(\mathcal{A},\mathcal{B})$. By Corollary 3.3, the sequences of symmetric conductances $\{c_N^s(\eta,\xi) : N \geq 1\}$, $(\eta,\xi) \in \mathbb{B}^s$, are ordered. It follows from this fact and from the proof of Lemmas 4.1 in Beltrán and Landim (2011) that there exists constants $0 < c_0 < C_0 < \infty$ such that

$$c_0 \;<\; \liminf_{N\to\infty} \frac{\mathrm{cap}_N^s(\mathcal{A},\mathcal{B})}{\boldsymbol{c}_N^s(\mathcal{A},\mathcal{B})} \;\leq\; \limsup_{N\to\infty} \frac{\mathrm{cap}_N^s(\mathcal{A},\mathcal{B})}{\boldsymbol{c}_N^s(\mathcal{A},\mathcal{B})} \;\leq\; C_0\,. \tag{5.4}$$

6. Proof of Theorem 2.1

In view of Theorem 5.1 in Landim (2015), Theorem 2.1 follows from from condition (H3) and from Propositions 6.1 below. Denote by $\psi_{\mathcal{E}} : \mathcal{E} \to \{1,\dots,\mathfrak{n}\}$ the projection defined by $\psi_{\mathcal{E}}(\eta) = x$ if $\eta \in \mathcal{E}_x$:

$$\psi_{\mathcal{E}}(\eta) \;=\; \sum_{x\in S} x\, \mathbf{1}\{\eta \in \mathcal{E}_x\}\,.$$

Proposition 6.1. *Fix $x \in S$ and a configuration $\eta \in \mathcal{E}_x$. Starting from η, the speeded-up, hidden Markov chain $X_N(t) = \psi_{\mathcal{E}}\big(\eta^{\mathcal{E}}(\theta_N t)\big)$ converges in the Skorohod topology to the continuous-time Markov chain $X_{\mathcal{E}}(t)$, introduced in Theorem 2.1, which starts from x.*

Lemma 6.2. *For every $x \in S$ for which $\mathcal{E}_x$ is not a singleton and for all $\eta \neq \xi \in \mathcal{E}_x$,*

$$\lim_{N\to\infty} \frac{\mathrm{cap}_N(\mathcal{E}_x, \breve{\mathcal{E}}_x)}{\mathrm{cap}_N(\eta,\xi)} \;=\; 0\,.$$

Proof: Fix $x \in S$. By (4.3), applied to $A = \mathcal{E}_x$, $B = \breve{\mathcal{E}}_x$, and by assumption (H1),

$$\lim_{N\to\infty} \theta_N \frac{\mathrm{cap}_N(\mathcal{E}_x, \breve{\mathcal{E}}_x)}{\mu_N(\mathcal{E}_x)} \;=\; \sum_{y\neq x} r_{\mathcal{E}}(x,y) \;\in\; \mathbb{R}_+\,.$$

The claim of the lemma follows from this equation, from assumption (H2) and from the fact that $\alpha_N/\theta_N \to 0$. □

Proof of Proposition 6.1: In view of Theorem 2.1 in Beltrán and Landim (2012), the claim of the proposition follows from condition (H1), and from Lemma 6.2. □

7. Proof of Theorem 2.7

The proof of Theorem 2.7 is divided in several steps.

1. The measure of the metastable sets. We start proving that condition (H0) is in force. Recall from Section 2 that we denote by $X_R(t)$ the E-valued chain which jumps from η to ξ at rate $R(\eta,\xi)$. Denote by $\mathcal{C}_1,\dots,\mathcal{C}_{\mathfrak{m}}$ the equivalent classes of the chain $X_R(t)$.

Assertion 7.A. For all $1 \leq j \leq \mathfrak{m}$, and for all $\eta \neq \xi \in \mathcal{C}_j$, there exists $m(\eta,\xi) \in (0,\infty)$ such that

$$\lim_{N\to\infty} \frac{\mu_N(\eta)}{\mu_N(\xi)} \;=\; m(\eta,\xi)\,.$$

Proof: Fix $1 \le j \le \mathfrak{m}$ and $\eta \neq \xi \in \mathcal{C}_j$. By assumption, there exists a path $(\eta = \eta_0, \dots, \eta_n = \xi)$ such that $R(\eta_i, \eta_{i+1}) > 0$ for $0 \le i < n$. On the other hand, since μ_N is an invariant probability measure,

$$\begin{aligned}\lambda_N(\xi)\,\mu_N(\xi) \;&=\; \sum_{\zeta_0,\zeta_1,\dots,\zeta_{n-1}\in E} \mu_N(\zeta_0)\,\lambda_N(\zeta_0)\,p_N(\zeta_0,\zeta_1)\cdots p_N(\zeta_{n-1},\xi)\\ &\ge\; \mu_N(\eta_0)\,\lambda_N(\eta_0)\,p_N(\eta_0,\eta_1)\cdots p_N(\eta_{n-1},\xi)\;.\end{aligned}$$

Therefore,

$$\frac{\mu_N(\xi)}{\mu_N(\eta)} \;\ge\; \frac{\lambda_N(\eta)}{\lambda_N(\xi)}\,p_N(\eta,\eta_1)\cdots p_N(\eta_{n-1},\xi)\;.$$

Since $R(\eta_i, \eta_{i+1}) > 0$ for $0 \le i < n$, by (2.5), $p_N(\eta_i, \eta_{i+1})$ converges to $p(\eta_i, \eta_{i+1}) > 0$. For the same reason, $\alpha_N \lambda_N(\eta)$ converges to $\lambda(\eta) \in (0, \infty)$. Finally, as ξ and η belong to the same equivalent class, there exists a path from ξ to η with similar properties to the one from η to ξ, so that $\alpha_N \lambda_N(\xi)$ converges to $\lambda(\xi) \in (0, \infty)$. In conclusion,

$$\liminf_{N\to\infty} \frac{\mu_N(\xi)}{\mu_N(\eta)} \;>\; 0\;.$$

Replacing η by ξ we obtain that $\liminf \mu_N(\eta)/\mu_N(\xi) > 0$. Since by Lemma 3.1 the sequences $\{\mu_N(\zeta) : N \ge 1\}$, $\zeta \in E$, are ordered, $\mu_N(\eta)/\mu_N(\xi)$ must converge to some value in $(0, \infty)$. □

By the previous assertion for every $x \in S$ and $\eta \in \mathcal{E}_x$,

$$m_x(\eta) \;:=\; \lim_{N\to\infty} \frac{\mu_N(\eta)}{\mu_N(\mathcal{E}_x)} \;\in\; (0,1]\;, \tag{7.1}$$

where we adopted the convention established in condition (H1) of Section 2.

2. The time-scale. In this subsection, we introduce a time-scale γ_N, we prove that it is much longer than α_N and that it is of the same order of θ_N. In particular the requirement $\alpha_N/\theta_N \to 0$ is in force.

Denote by $\{\eta_t^{\mathcal{E}} : t \ge 0\}$ the trace of η_t^N on the set $\mathcal{E}$, and by $R_N^{\mathcal{E}} : \mathcal{E} \times \mathcal{E} \to \mathbb{R}_+$ the jump rates of $\eta_t^{\mathcal{E}}$. Let

$$\frac{1}{\gamma_N} \;=\; \sum_{x\in S} \sum_{\eta\in\mathcal{E}_x} \sum_{\xi\in\breve{\mathcal{E}}_x} R_N^{\mathcal{E}}(\eta,\xi)\;, \tag{7.2}$$

where $\breve{\mathcal{E}}_x$ has been introduced in (2.4). The sequence γ_N represents the time needed to reach the set $\breve{\mathcal{E}}_x$ starting from $\mathcal{E}_x$ for some $x \in S$. This time scale might be longer for other sets $\mathcal{E}_y$, $y \neq x$, but it is of the order γ_N at least for one $x \in S$. We could as well have defined γ_N as $\max_{x\in S} \max_{\eta\in\mathcal{E}_x} \max_{\xi\in\breve{\mathcal{E}}_x} R_N^{\mathcal{E}}(\eta,\xi)$.

Assertion 7.B. The time scale γ_N is much longer than the time-scale α_N:

$$\lim_{N\to\infty} \frac{\alpha_N}{\gamma_N} \;=\; 0\;.$$

Proof: We have to show that $\alpha_N R_N^{\mathcal{E}}(\eta,\xi)$ converges to 0 as $N \uparrow \infty$, for all $\eta \in \mathcal{E}_x$, $\xi \in \mathcal{E}_y$, $x \neq y \in S$. Fix $x \neq y \in S$, $\eta \in \mathcal{E}_x$, $\xi \in \mathcal{E}_y$. Since $\mathcal{E}_x$ is a recurrent class, $R(\eta, \zeta) = 0$ for all $\zeta \notin \mathcal{E}_x$. On the other hand, by Beltrán and Landim (2010, Proposition 6.1) and by the strong Markov property,

$$R_N^{\mathcal{E}}(\eta,\xi) \;=\; \lambda_N(\eta)\,\mathbb{P}_\eta[H_\xi = H_{\mathcal{E}}^+] \;=\; R_N(\eta,\xi) \;+\; \sum_{\zeta\notin\mathcal{E}} R_N(\eta,\zeta)\,\mathbb{P}_\zeta[H_\xi = H_{\mathcal{E}}]\;.$$

Since $R(\eta,\zeta)=0$ for all $\zeta\notin\mathcal{E}_x$, it follows from the previous identity and from the definition of $R(\eta,\zeta)$ that $\alpha_N R_N^{\mathcal{E}}(\eta,\xi)\to 0$, as claimed. □

By Assertion 3.A, for all $x\in S$, $\eta\in\mathcal{E}_x$, $\xi\in\breve{\mathcal{E}}_x$, with the convention adopted in condition (H1) of Section 2,

$$r_{\mathcal{E}}(\eta,\xi) \;:=\; \lim_{N\to\infty}\gamma_N\, R_N^{\mathcal{E}}(\eta,\xi) \;\in\; [0,1]\,. \tag{7.3}$$

Assertion 7.C. For all $x\in S$,

$$\ell_x \;:=\; \lim_{N\to\infty}\gamma_N\,\frac{\mathrm{cap}_N(\mathcal{E}_x,\breve{\mathcal{E}}_x)}{\mu_N(\mathcal{E}_x)} \;\in\; \mathbb{R}_+\,. \quad \text{Moreover}\,, \;\; \ell \;=\; \sum_{x\in S}\ell_x \;>\; 0\,.$$

Proof: By (4.3), applied to $A=\mathcal{E}_x$, $B=\breve{\mathcal{E}}_x$, by (7.1) and by (7.3),

$$\lim_{N\to\infty}\gamma_N\,\frac{\mathrm{cap}_N(\mathcal{E}_x,\breve{\mathcal{E}}_x)}{\mu_N(\mathcal{E}_x)} \;=\; \sum_{\eta\in\mathcal{E}_x} m_x(\eta)\sum_{\xi\in\breve{\mathcal{E}}_x} r_{\mathcal{E}}(\eta,\xi) \;\in\; \mathbb{R}_+\,,$$

which completes the proof of the first claim of the assertion.

By (7.2) and by definition of $r_{\mathcal{E}}(\eta,\xi)$,

$$\sum_{x\in S}\sum_{\eta\in\mathcal{E}_x}\sum_{\xi\in\breve{\mathcal{E}}_x} r_{\mathcal{E}}(\eta,\xi) \;=\; 1\,,$$

so that

$$\ell \;=\; \sum_{x\in S}\ell_x \;=\; \sum_{x\in S}\sum_{\eta\in\mathcal{E}_x} m_x(\eta)\sum_{\xi\in\breve{\mathcal{E}}_x} r_{\mathcal{E}}(\eta,\xi) \;\geq\; \min_{x\in S}\min_{\eta\in\mathcal{E}_x} m_x(\eta) \;>\; 0\,,$$

which is the second claim of the assertion. □

It follows from Assertion 7.C that the time-scale γ_N is of the same order of θ_N in the sense that γ_N/θ_N converges as $N\uparrow\infty$:

$$\lim_{N\to\infty}\frac{\gamma_N}{\theta_N} \;=\; \ell \;\in\; (0,\infty)\,. \tag{7.4}$$

3. The average jump rate, condition (H1). Denote by $r_N(\mathcal{E}_x,\mathcal{E}_y)$ the mean rate at which the trace process jumps from $\mathcal{E}_x$ to $\mathcal{E}_y$:

$$r_N(\mathcal{E}_x,\mathcal{E}_y) \;=\; \frac{1}{\mu_N(\mathcal{E}_x)}\sum_{\eta\in\mathcal{E}_x}\mu_N(\eta)\sum_{\xi\in\mathcal{E}_y} R_N^{\mathcal{E}}(\eta,\xi)\,. \tag{7.5}$$

Next lemma follows from (7.1), (7.3) and (7.4).

Lemma 7.1. *For every $x\neq y\in S$,*

$$r_{\mathcal{E}}(x,y) \;:=\; \lim_{N\to\infty}\theta_N\, r_N(\mathcal{E}_x,\mathcal{E}_y) \;=\; \frac{1}{\ell}\sum_{\eta\in\mathcal{E}_x} m_x(\eta)\sum_{\xi\in\mathcal{E}_y} r_{\mathcal{E}}(\eta,\xi) \;\in\; \mathbb{R}_+$$

4. Inside the metastable sets, condition (H2). Next assertion shows that condition (H2) is in force.

Assertion 7.D. For every $x\in S$ for which $\mathcal{E}_x$ is not a singleton and for all $\eta\neq\xi\in\mathcal{E}_x$, there exist constants $0<c_0<C_0<\infty$ such that

$$c_0 \;\leq\; \liminf_{N\to\infty}\alpha_N\,\frac{\mathrm{cap}_N(\eta,\xi)}{\mu_N(\mathcal{E}_x)} \;\leq\; \limsup_{N\to\infty}\alpha_N\,\frac{\mathrm{cap}_N(\eta,\xi)}{\mu_N(\mathcal{E}_x)} \;\leq\; C_0\,.$$

Proof: Fix $x \in S$ for which $\mathcal{E}_x$ is not a singleton, and $\eta \neq \xi \in \mathcal{E}_x$. On the one hand, by definition of the capacity

$$\alpha_N \, \frac{\mathrm{cap}_N(\eta,\xi)}{\mu_N(\mathcal{E}_x)} \;\le\; \frac{\mu_N(\eta)}{\mu_N(\mathcal{E}_x)}\, \alpha_N\, \lambda_N(\eta)\;.$$

By (2.5) and (7.1), the right hand side converges to $\lambda(\eta) m_x(\eta) < \infty$, which proves one of the inequalities.

On the other hand, as $\mathcal{E}_x$ is an equivalent class which is not a singleton, $\lambda(\zeta) > 0$ for all $\zeta \in \mathcal{E}_x$, or, in other words, $\mathcal{E}_x \subset E_0$. Since $\eta \sim \xi$, there exists a path $(\eta = \eta_0, \dots, \eta_n = \xi)$ such that $R(\eta_i, \eta_{i+1}) > 0$ for $0 \le i < n$. Since,

$$\mathbb{P}_\eta\big[H_\xi < H_\eta^+\big] \;\ge\; p_N(\eta,\eta_1)\cdots p_N(\eta_{n-1},\xi)\;,$$

in view of the formula (2.2) for the capacity, we have that

$$\alpha_N \, \frac{\mathrm{cap}_N(\eta,\xi)}{\mu_N(\mathcal{E}_x)} \;\ge\; \frac{\mu_N(\eta)}{\mu_N(\mathcal{E}_x)}\, \alpha_N\, \lambda_N(\eta)\, p_N(\eta,\eta_1)\cdots p_N(\eta_{n-1},\xi)\;.$$

The right hand side converges to $m_x(\eta)\lambda(\eta)p(\eta,\eta_1)\cdots p(\eta_{n-1},\xi) > 0$, which completes the proof of the assertion. □

5. Condition (H3) holds. To complete the proof of Theorem 2.7 it remains to show that the chain η_t^N spends a negligible amount of time on the set Δ in the time scale θ_N.

Lemma 7.2. *For every $t > 0$,*

$$\lim_{N\to\infty} \max_{\eta\in E} \mathbb{E}_\eta\Big[\int_0^t \mathbf{1}\{\eta^N_{s\theta_N} \in \Delta\}\, ds\Big] \;=\; 0\;.$$

Proof: Since $\alpha_N/\theta_N \to 0$, a change of variables in the time integral and the Markov property show that for every $\eta \in E$, for every $T > 0$ and for every N large enough,

$$\mathbb{E}_\eta\Big[\int_0^t \mathbf{1}\{\eta^N_{s\theta_N} \in \Delta\}\, ds\Big] \;\le\; \frac{2t}{T} \max_{\xi\in E} \mathbb{E}_\xi\Big[\int_0^T \mathbf{1}\{\eta^N_{s\alpha_N} \in \Delta\}\, ds\Big]\;.$$

Note that the process on the right hand side is speeded up by α_N instead of θ_N.

We estimate the expression on the right hand side of the previous formula. We may, of course, restrict the maximum to Δ. Let T_1 be the first time the chain η_t^N hits $\mathcal{E}$ and let T_2 be the time it takes for the process to return to Δ after T_1:

$$T_1 \;=\; H_{\mathcal{E}}\;, \quad T_2 \;=\; \inf\big\{s > 0 : \eta^N_{T_1+s} \in \Delta\big\}\;.$$

Fix $\eta \in \Delta$ and note that

$$\begin{aligned} &\mathbb{E}_\eta\Big[\frac{1}{T}\int_0^T \mathbf{1}\{\eta^N_{s\alpha_N} \in \Delta\}\, ds\Big] \\ &\qquad\le\; \mathbb{P}_\eta\big[T_1 > t_0\alpha_N\big] \;+\; \mathbb{P}_\eta\big[T_2 < T\alpha_N\big] \;+\; \frac{t_0}{T} \end{aligned} \tag{7.6}$$

for all $t_0 > 0$ because the time average is bounded by 1 and because on the set $\{T_1 \le t_0\alpha_N\} \cap \{T_2 \ge T\alpha_N\}$ the time average is bounded by t_0/T. By Assertion 7.E below, the first term on the right hand side vanishes as $N \uparrow \infty$ and then $t_0 \uparrow \infty$. On the other hand, by the strong Markov property, the second term is bounded by $\max_{\xi\in\mathcal{E}} \mathbb{P}_\xi[H_\Delta \le T\alpha_N]$. By definition of the set $\mathcal{E}$, for every $\eta \in \mathcal{E}$ and every $\xi \in \Delta$, $\alpha_N R_N(\eta,\xi) \to 0$ as $N \uparrow \infty$. This shows that for every $T > 0$ the second term on the right hand side of (7.6) vanishes as $N \uparrow \infty$, which completes the proof of the lemma. □

Assertion 7.E. For every $\eta \in \Delta$,

$$\lim_{t\to\infty} \limsup_{N\to\infty} \mathbb{P}_\eta\big[H_{\mathcal{E}} \ge t\alpha_N\big] \;=\; 0\,.$$

Proof: Recall that we denote by $X_R(t)$ the continuous-time Markov chain on E which jumps from η to ξ at rate $R(\eta,\xi) = \lim_N \alpha_N R_N(\eta,\xi)$. Note that the set $\mathcal{E}$ consists of recurrent points for the chain $X_R(t)$, while points in Δ are transient. Since the jump rates converge, the chain $\eta^N_{t\alpha_N}$ converges in the Skorohod topology to $X_R(t)$. Therefore, for all $t>0$, $\eta\in\Delta$,

$$\limsup_{N\to\infty} \mathbb{P}_\eta\big[H_{\mathcal{E}} \ge t\,\alpha_N\big] \;\le\; \mathbf{P}_\eta\big[H_{\mathcal{E}} \ge t\big]\,,$$

where $\mathbf{P}_\eta$ stands for the law of the chain $X_R(t)$ starting from η. Since the set of recurrent points for $X_R(t)$ is equal to $\mathcal{E} = \Delta^c$, the previous probability vanishes as $t\uparrow\infty$. □

We conclude this section with an observation concerning the capacities of the metastable sets $\mathcal{E}_x$.

Assertion 7.F. The sequences $\{\mathrm{cap}_N(\mathcal{E}_x, \breve{\mathcal{E}}_x)/\mu_N(\mathcal{E}_x) : N\ge 1\}$, $x\in S$, are ordered.

Proof: Fix $x\in S$. By (4.3) applied to $A=\mathcal{E}_x$, $B=\breve{\mathcal{E}}_x$,

$$\mathrm{cap}_N(\mathcal{E}_x, \breve{\mathcal{E}}_x) \;=\; \sum_{\eta\in\mathcal{E}_x} \mu_N(\eta) \sum_{\xi\in\breve{\mathcal{E}}_x} R^{\mathcal{E}}_N(\eta,\xi)\,.$$

The claim of the assertion follows from this identity, from Assertion 3.A and from (7.1). □

8. Proof of Theorem 2.12

Theorem 2.12 is proved in several steps.

1. The measure of configurations in $\mathcal{G}_a$. We assumed in (H0) that all configurations in a set $\mathcal{F}_x$ have measure of the same order. We prove below in Assertion 8.A that a similar property holds for the sets $\mathcal{G}_a$.

Let

$$\lambda^{\mathcal{F}}_N(\mathcal{F}_x) \;=\; \sum_{y:y\neq x} r^{\mathcal{F}}_N(\mathcal{F}_x,\mathcal{F}_y)\,,\quad p^{\mathcal{F}}_N(\mathcal{F}_x,\mathcal{F}_y) \;=\; \frac{r^{\mathcal{F}}_N(\mathcal{F}_x,\mathcal{F}_y)}{\lambda^{\mathcal{F}}_N(\mathcal{F}_x)} \;\text{ if }\; \lambda^{\mathcal{F}}_N(\mathcal{F}_x)>0\,.$$

Denote by P_0 the subset of points in P such that $\lambda_{\mathcal{F}}(x) = \sum_{y\neq x} r_{\mathcal{F}}(x,y) > 0$. For all $x\in P_0$ let $p_{\mathcal{F}}(x,y) = r_{\mathcal{F}}(x,y)/\lambda_{\mathcal{F}}(x)$. It follows from assumption (H1) that for all x, z in P, $y\in P_0$,

$$\lim_{N\to\infty} \beta_N\, \lambda^{\mathcal{F}}_N(\mathcal{F}_x) \;=\lambda_{\mathcal{F}}(x)\,,\quad \lim_{N\to\infty} p^{\mathcal{F}}_N(\mathcal{F}_y,\mathcal{F}_z) \;=\; p_{\mathcal{F}}(y,z)\,. \tag{8.1}$$

Recall that $X_{\mathcal{F}}(t)$ is the P-valued Markov chain which jumps from x to y at rate $r_{\mathcal{F}}(x,y)$. Denote by C_a, $a\in P_1 = \{1,\dots,\mathfrak{q}_1\}$, the equivalent classes of the Markov chain $X_{\mathcal{F}}(t)$, and let $\mathcal{C}_a = \cup_{x\in C_a}\mathcal{F}_x$. All configurations in a set $\mathcal{C}_a$ have probability of the same order.

Assertion 8.A. For all equivalent classes C_a, $a\in P_1$, and for all $\eta\neq\xi\in\mathcal{C}_a$, there exists $m(\eta,\xi)\in(0,\infty)$ such that

$$\lim_{N\to\infty} \frac{\mu_N(\eta)}{\mu_N(\xi)} \;=\; m(\eta,\xi)\,.$$

Proof: The argument is very close to the one of Assertion 7.A Denote by $\bar{X}_N(t)$ the chain $\eta^{\mathcal{F}}(t)$ in which each set $\mathcal{F}_x$ has been collapsed to a point. We refer to the Section 3 of Gaudillière and Landim (2014) for a precise definition of the collapsed chain and for the proof of the results used below.

The chain $\bar{X}_N(t)$ takes value in the set P, its jump rate from x to y, denoted by $\bar{r}_N(x,y)$, is equal to $r_N^{\mathcal{F}}(\mathcal{F}_x,\mathcal{F}_y)$ introduced in (2.3), and its unique invariant probability measure, denoted by $\bar{\mu}_N(x)$, is given by $\bar{\mu}_N(x) = \mu_N(\mathcal{F}_x)/\mu_N(\mathcal{F})$.

Fix an equivalent class C_a and $\eta \neq \xi \in \mathcal{C}_a$. If η and ξ belong to the same set $\mathcal{F}_x$, the claim follows from Assumption (H0). Suppose that $\eta \in \mathcal{F}_x$, $\xi \in \mathcal{F}_y$ for some $x \neq y \in C_a$. By assumption, there exists a path $(x = x_0, \dots, x_n = y)$ such that $r_{\mathcal{F}}(x_i, x_{i+1}) > 0$ for $0 \le i < n$.

Denote by $\bar{\lambda}_N(x)$, $x \in P$, the holding rates of the collapsed chain $\bar{X}_N(t)$, and by $\bar{p}_N(x,y)$, $x \neq y \in P$, the jump probabilities. Since $\bar{\mu}_N$ is the invariant probability measure for the collapsed chain,

$$\begin{aligned}\bar{\lambda}_N(y)\,\bar{\mu}_N(y) \;&=\; \sum_{z_0,z_1,\dots,z_{n-1}\in P} \bar{\mu}_N(z_0)\,\bar{\lambda}_N(z_0)\,\bar{p}_N(z_0,z_1)\cdots\bar{p}_N(z_{n-1},y)\\ &\ge\; \bar{\mu}_N(x_0)\,\bar{\lambda}_N(x_0)\,\bar{p}_N(x_0,x_1)\cdots\bar{p}_N(x_{n-1},y)\;.\end{aligned}$$

Therefore,

$$\frac{\bar{\mu}_N(y)}{\bar{\mu}_N(x)} \;\ge\; \frac{\bar{\lambda}_N(x)}{\bar{\lambda}_N(y)}\,\bar{p}_N(x,x_1)\cdots\bar{p}_N(x_{n-1},y)\;.$$

Since $r_{\mathcal{F}}(x_i,x_{i+1}) > 0$ for $0 \le i < n$, by (8.1), $\bar{p}_N(x_i,x_{i+1})$ converges to $p_{\mathcal{F}}(x_i,x_{i+1}) > 0$. For the same reason, $\beta_N\bar{\lambda}_N(x) = \beta_N\lambda_N^{\mathcal{F}}(\mathcal{F}_x)$ converges to $\lambda_{\mathcal{F}}(x) \in (0,\infty)$. As y and x share the same properties, inverting their role we obtain that $\beta_N\bar{\lambda}_N(y)$ converges to $\lambda_{\mathcal{F}}(y) \in (0,\infty)$. In conclusion,

$$\liminf_{N\to\infty} \frac{\bar{\mu}_N(x)}{\bar{\mu}_N(y)} \;>\; 0\;.$$

Replacing x by y we obtain that $\liminf \bar{\mu}_N(y)/\bar{\mu}_N(x) > 0$. By Gaudillière and Landim (2014), $\bar{\mu}_N(z) = \mu_N(\mathcal{F}_z)$, $z \in P$. To complete the proof it remains to recall the statement of Lemma 3.1 and Assumption (H0). □

By the previous assertion for every $a \in Q$ and $\eta \in \mathcal{G}_a$,

$$m_a^*(\eta) \;:=\; \lim_{N\to\infty} \frac{\mu_N(\eta)}{\mu_N(\mathcal{G}_a)} \;\in\; (0,1]\;. \tag{8.2}$$

Thus, assumption (H0) holds for the partition $\{\mathcal{G}_1,\dots,\mathcal{G}_{\mathfrak{q}},\Delta_{\mathcal{G}}\}$.

2. The time scale. We prove in this subsection that the time-scale β_N^+ introduced in (2.10) is much longer than β_N.

Assertion 8.B. We have that

$$\lim_{N\to\infty} \frac{\beta_N}{\beta_N^+} \;=\; 0\;.$$

Proof: We have to show that

$$\lim_{N\to\infty} \beta_N \frac{\mathrm{cap}_N(\mathcal{G}_a,\breve{\mathcal{G}}_a)}{\mu_N(\mathcal{G}_a)} \;=\; 0$$

for each $a \in Q$. Fix $a \in Q$ and recall from (2.8) the definition of the set $\mathcal{G}_a$. Since G_a is a recurrent class for the chain $X_{\mathcal{F}}(t)$, $r_{\mathcal{F}}(x,y) = 0$ for all $x \in G_a$, $y \in P \setminus G_a$. By definition of the capacity,

$$\begin{aligned}\frac{\mathrm{cap}_N(\mathcal{G}_a, \breve{\mathcal{G}}_a)}{\mu_N(\mathcal{G}_a)} &= \sum_{\eta\in\mathcal{G}_a} \frac{\mu_N(\eta)}{\mu_N(\mathcal{G}_a)} \lambda_N(\eta)\, \mathbb{P}_\eta\big[H_{\breve{\mathcal{G}}_a} < H^+_{\mathcal{G}_a}\big] \\ &\le \sum_{\eta\in\mathcal{G}_a} \frac{\mu_N(\eta)}{\mu_N(\mathcal{G}_a)} \lambda_N(\eta)\, \mathbb{P}_\eta\big[H_{\mathcal{F}\setminus\mathcal{G}_a} < H^+_{\mathcal{G}_a}\big]\,.\end{aligned}$$

By Beltrán and Landim (2010, Proposition 6.1), this sum is equal to

$$\sum_{\eta\in\mathcal{G}_a} \frac{\mu_N(\eta)}{\mu_N(\mathcal{G}_a)} \sum_{\xi\in\mathcal{F}\setminus\mathcal{G}_a} R^{\mathcal{F}}_N(\eta,\xi) = \sum_{x\in G_a} \frac{\mu_N(\mathcal{F}_x)}{\mu_N(\mathcal{G}_a)} \sum_{y\in P\setminus G_a} r^{\mathcal{F}}_N(x,y)\,.$$

Since $r_{\mathcal{F}}(x,y) = 0$ for all $x \in G_a$, $y \in P \setminus G_a$, by assumption (H1) the previous sum multiplied by β_N converges to 0 as $N \uparrow \infty$. □

3. Condition (H1) is fulfilled by the partition $\{\mathcal{G}_1, \dots, \mathcal{G}_{\mathfrak{q}}, \Delta_{\mathcal{G}}\}$. We first obtain an alternative formula for the time-scale β^+_N. The arguments and the ideas are very similar to the ones presented in the previous section. Let

$$\frac{1}{\gamma_N} = \sum_{a\in Q} \sum_{\eta\in\mathcal{G}_a} \sum_{\xi\in\breve{\mathcal{G}}_a} R^{\mathcal{G}}_N(\eta,\xi)\,.$$

By Assertion 3.A, for all $a \in Q$, $\eta \in \mathcal{G}_a$, $\xi \in \breve{\mathcal{G}}_a$, with the convention adopted in condition (H1) of Section 2,

$$r_{\mathcal{G}}(\eta,\xi) := \lim_{N\to\infty} \gamma_N\, R^{\mathcal{G}}_N(\eta,\xi) \in [0,1]\,. \tag{8.3}$$

Assertion 8.C. For all $a \in Q$,

$$\hat\lambda_{\mathcal{G}}(a) := \lim_{N\to\infty} \gamma_N \frac{\mathrm{cap}_N(\mathcal{G}_a, \breve{\mathcal{G}}_a)}{\mu_N(\mathcal{G}_a)} \in \mathbb{R}_+\,. \quad \text{Moreover}\,, \quad \hat\lambda_{\mathcal{G}} = \sum_{a\in Q} \hat\lambda_{\mathcal{G}}(a) > 0\,.$$

Proof: Fix $a \in Q$. By (4.3), applied to $A = \mathcal{G}_a$, $B = \breve{\mathcal{G}}_a$, by (8.2) and by (8.3),

$$\lim_{N\to\infty} \gamma_N \frac{\mathrm{cap}_N(\mathcal{G}_a, \breve{\mathcal{G}}_a)}{\mu_N(\mathcal{G}_a)} = \sum_{\eta\in\mathcal{G}_a} m^*_a(\eta) \sum_{\xi\in\breve{\mathcal{G}}_a} r_{\mathcal{G}}(\eta,\xi) \in \mathbb{R}_+\,,$$

which completes the proof of the first claim of the assertion.

By definition of γ_N and by definition of $r_{\mathcal{G}}(\eta,\xi)$,

$$\sum_{a\in Q} \sum_{\eta\in\mathcal{G}_a} \sum_{\xi\in\breve{\mathcal{G}}_a} r_{\mathcal{G}}(\eta,\xi) = 1\,,$$

so that

$$\hat\lambda_{\mathcal{G}} = \sum_{a\in Q} \hat\lambda_{\mathcal{G}}(a) = \sum_{a\in Q} \sum_{\eta\in\mathcal{G}_a} m^*_a(\eta) \sum_{\xi\in\breve{\mathcal{G}}_a} r_{\mathcal{G}}(\eta,\xi) \ge \min_{a\in Q} \min_{\eta\in\mathcal{G}_a} m^*_a(\eta) > 0\,,$$

which is the second claim of the assertion. □

It follows from the previous assertion that the time-scale γ_N is of the same order of β_N^+:

$$\lim_{N\to\infty} \frac{\gamma_N}{\beta_N^+} \;=\; \hat\lambda_{\mathcal{G}} \;\in (0,\infty)\,. \tag{8.4}$$

Denote by $r_N^{\mathcal{G}}(\mathcal{G}_a,\mathcal{G}_b)$ the mean rate at which the trace process jumps from $\mathcal{G}_a$ to $\mathcal{G}_b$:

$$r_N^{\mathcal{G}}(\mathcal{G}_a,\mathcal{G}_b) \;:=\; \frac{1}{\mu_N(\mathcal{G}_a)} \sum_{\eta\in\mathcal{G}_a} \mu_N(\eta) \sum_{\xi\in\mathcal{G}_b} R_N^{\mathcal{G}}(\eta,\xi)\,. \tag{8.5}$$

Lemma 8.1. *For every $a\neq b\in Q$,*

$$r_{\mathcal{G}}(a,b) \;:=\; \lim_{N\to\infty} \beta_N^+\, r_N^{\mathcal{G}}(\mathcal{G}_a,\mathcal{G}_b) \;=\; \frac{1}{\hat\lambda_{\mathcal{G}}} \sum_{\eta\in\mathcal{G}_a} m_a^*(\eta) \sum_{\xi\in\mathcal{G}_b} r_{\mathcal{G}}(\eta,\xi) \;\in\; \mathbb{R}_+$$

Moreover,

$$\sum_{a\in Q} \sum_{b:b\neq a} r_{\mathcal{G}}(a,b) \;=\; 1\,.$$

Proof: The first claim of this lemma follows from (8.2), (8.3) and (8.4). On the other hand, by the explicit formula for $r_{\mathcal{G}}(a,b)$ and by the formula for $\hat\lambda_{\mathcal{G}}(a)$ obtained in the previous assertion,

$$\sum_{a\in Q} \sum_{b:b\neq a} r_{\mathcal{G}}(a,b) \;=\; \frac{1}{\hat\lambda_{\mathcal{G}}} \sum_{a\in Q} \sum_{\eta\in\mathcal{G}_a} m_a^*(\eta) \sum_{b:b\neq a} \sum_{\xi\in\mathcal{G}_b} r_{\mathcal{G}}(\eta,\xi) \;=\; \frac{1}{\hat\lambda_{\mathcal{G}}} \sum_{a\in Q} \hat\lambda_{\mathcal{G}}(a)\,.$$

This expression is equal to 1 by definition of $\hat\lambda_{\mathcal{G}}$. □

4. Condition (H2) is fulfilled by the partition $\{\mathcal{G}_1,\dots,\mathcal{G}_{\mathfrak{q}},\Delta_{\mathcal{G}}\}$. The proof of condition (H2) is based on the next assertion.

Assertion 8.D. For every $a\in Q$ for which $\mathcal{G}_a$ is not a singleton and for all $\eta\neq\xi\in\mathcal{G}_a$,

$$\liminf_{N\to\infty} \beta_N \frac{\mathrm{cap}_N(\eta,\xi)}{\mu_N(\mathcal{G}_a)} \;>\; 0\,.$$

Proof: Throughout this proof c_0 represents a positive real number independent of N and which may change from line to line. Fix $a\in Q$ for which $\mathcal{G}_a$ is not a singleton, and $\eta\neq\xi\in\mathcal{G}_a$. By definition, $\mathcal{G}_a=\cup_{x\in G_a}\mathcal{F}_x$. If η and ξ belongs to the same $\mathcal{F}_x$, the result follows from assumption (H2) and from Assertion 8.A.

Fix $\eta\in\mathcal{F}_x$ and $\xi\in\mathcal{F}_y$ for some $x\neq y$, $\mathcal{F}_x\cup\mathcal{F}_y\subset\mathcal{G}_a$. Recall that we denote by $\mathrm{cap}_N^{\mathrm{s}}(\mathcal{A},\mathcal{B})$ the capacity between two disjoint subsets $\mathcal{A}$, $\mathcal{B}$ of E with respect to the reversible chain introduced in (4.2).

Since G_a is a recurrent class for the chain $X_{\mathcal{F}}(t)$, there exists a sequence $(x=x_0,x_1,\dots,x_n=y)$ such that $r_{\mathcal{F}}(x_i,x_{i+1})>0$ for $0\le i<n$. In view of assumptions (H0) and (H1), there exist $\xi_i\in\mathcal{F}_{x_i}$, $\eta_{i+1}\in\mathcal{F}_{x_{i+1}}$ such that $\beta_N R_N^{\mathcal{F}}(\xi_i,\eta_{i+1})\ge c_0$. Therefore, by Corollary 4.4 and (4.4),

$$\beta_N\,\mathrm{cap}_N^{\mathrm{s}}(\xi_i,\eta_{i+1}) \;\ge\; \frac{\beta_N}{2|E|}\,\mathrm{cap}_N(\xi_i,\eta_{i+1}) \;\ge\; c_0\,\mu_N(\xi_i)\,, \tag{8.6}$$

so that, by (5.4), $\beta_N\,\boldsymbol{c}_N^{\mathrm{s}}(\xi_i,\eta_{i+1})\ge c_0\,\mu_N(\xi_i)$.

Since the configuration η and ξ_0 belongs to the same set $\mathcal{F}_x$, by assumption (H2), $\beta_N^-\mathrm{cap}_N(\eta,\xi_0)/\mu_N(\mathcal{F}_x)\ge c_0$. A similar assertion holds for the pair of configurations

η_i, ξ_i, $1 \le i < n$, and for the pair η_n, ξ. Hence, if we set $\eta_0 = \eta$, $\xi_n = \xi$, by Corollary 4.4 and (5.4), we have that

$$\beta_N^- \boldsymbol{c}_N^s(\eta_i, \xi_i) \;\ge\; c_0 \mu_N(\mathcal{F}_{x_i}) \,.$$

By (8.2), we may replace $\mu_N(\mathcal{F}_{x_i})$ by $\mu_N(\mathcal{G}_a)$ in the previous inequality, and $\mu_N(\xi_i)$ by $\mu_N(\mathcal{G}_a)$ in (8.6). By (5.3),

$$\boldsymbol{c}_N^s(\eta, \xi) \;\ge\; \min_{0 \le i < n} \min \big\{ \boldsymbol{c}_N^s(\eta_i, \xi_i), \boldsymbol{c}_N^s(\xi_i, \eta_{i+1}), \boldsymbol{c}_N^s(\eta_n, \xi_n) \big\} \,.$$

Since $\beta_N^- \ll \beta_N$, it follows from the previous estimates that $\beta_N \boldsymbol{c}_N^s(\eta, \xi) \ge c_0 \mu_N(\mathcal{G}_a)$. To complete the proof, it remains to recall that, by Corollary 4.4 and (5.4), $\mathrm{cap}_N(\eta, \xi) \ge \mathrm{cap}_N^s(\eta, \xi) \ge c_0 \boldsymbol{c}_N^s(\eta, \xi)$. □

5. Condition (H3) is fulfilled by the partition $\{\mathcal{G}_1, \dots, \mathcal{G}_{\mathfrak{q}}, \Delta_{\mathcal{G}}\}$. Lemma 8.2 shows that it is enough to prove condition (H3) for the trace process $\eta^{\mathcal{F}}(t)$.

Lemma 8.2. *Assume that*

$$\lim_{N \to \infty} \max_{\eta \in \mathcal{F}} \mathbb{E}_\eta \Big[\int_0^t \mathbf{1}\{\eta^{\mathcal{F}}_{s\beta_N^+} \in \Delta_*\} \, ds \Big] \;=\; 0 \,,$$

where $\Delta_* = \cup_{x \in G_{\mathfrak{q}+1}} \mathcal{F}_x$ has been introduced in (2.8). Then,

$$\lim_{N \to \infty} \max_{\eta \in E} \mathbb{E}_\eta \Big[\int_0^t \mathbf{1}\{\eta^{N}_{s\beta_N^+} \in \Delta_{\mathcal{G}}\} \, ds \Big] \;=\; 0 \,.$$

Proof: Fix $\eta \in E$. Since $\Delta_{\mathcal{G}} = \Delta_* \cup \Delta_{\mathcal{F}}$,

$$\begin{aligned} &\mathbb{E}_\eta \Big[\int_0^t \mathbf{1}\{\eta_{s\beta_N^+} \in \Delta_{\mathcal{F}} \cup \Delta_*\} \, ds \Big] \\ &\quad \le\; \mathbb{E}_\eta \Big[\int_0^t \mathbf{1}\{\eta_{s\beta_N^+} \in \Delta_{\mathcal{F}}\} \, ds \Big] \;+\; \max_{\xi \in \mathcal{F}} \mathbb{E}_\xi \Big[\int_0^t \mathbf{1}\{\eta^{\mathcal{F}}_{s\beta_N^+} \in \Delta_*\} \, ds \Big] \,. \end{aligned}$$

The second term vanishes as $N \uparrow \infty$ by assumption. The first one is bounded by

$$\frac{\beta_N}{\beta_N^+} \sum_{n=0}^{[\beta_N^+/\beta_N]} \mathbb{E}_\eta \Big[\int_{nt}^{(n+1)t} \mathbf{1}\{\eta_{s\beta_N} \in \Delta_{\mathcal{F}}\} \, ds \Big] \,,$$

where $[r]$ stands for the integer part of r. By the Markov property, this expression is bounded above by

$$2 \max_{\xi \in E} \mathbb{E}_\xi \Big[\int_0^t \mathbf{1}\{\eta_{s\beta_N} \in \Delta_{\mathcal{F}}\} \, ds \Big] \,,$$

which vanishes as $N \uparrow \infty$ by assumption (H3). □

To prove that condition (H3) is fulfilled by the partition $\{\mathcal{G}_1, \dots, \mathcal{G}_{\mathfrak{q}}, \Delta_{\mathcal{G}}\}$ it remains to show that the assumption of the previous lemma is in force. The proof of this claim relies on the next assertion. Denote by $\mathbb{P}_\eta^{\mathcal{F}}$ the probability measure on $D(\mathbb{R}_+, \mathcal{F})$ induced by the trace chain $\eta_t^{\mathcal{F}}$ starting from η.

Assertion 8.E. For every $\eta \in \Delta_*$,

$$\lim_{t \to \infty} \limsup_{N \to \infty} \mathbb{P}_\eta^{\mathcal{F}} \big[H_{\mathcal{G}} \ge t\beta_N \big] \;=\; 0 \,.$$

Proof: Fix $\eta \in \mathcal{F}_x \subset \Delta_*$. Since the partition $\mathcal{F}_1, \dots, \mathcal{F}_{\mathfrak{p}}$, $\Delta_{\mathcal{F}}$ satisfy the conditions (H1)–(H3), by Proposition 6.1, starting from η the process $X_N(t) = \psi_{\mathcal{F}}(\eta^{\mathcal{F}}_{t\beta_N})$ converges in the Skorohod topology to the Markov chain $X_{\mathcal{F}}(t)$ on $P = \{1, \dots, \mathfrak{p}\}$ which starts from x and which jumps from y to z at rate $r_{\mathcal{F}}(y, z)$. Therefore,

$$\limsup_{N\to\infty} \mathbb{P}^{\mathcal{F}}_\eta \big[H_{\mathcal{G}} \ge t\,\beta_N \big] \;\le\; \mathbf{P}_x \big[H_R \ge t \big] \,,$$

where $\mathbf{P}_x$ represents the distribution of the chain $X_{\mathcal{F}}(t)$ starting from x and $R = \cup_{1\le a\le \mathfrak{q}} G_a$. Since R corresponds to the set of recurrent points of the chain $X_{\mathcal{F}}(t)$, the previous expression vanishes as $t \uparrow \infty$. □

Lemma 8.3. *For all $t > 0$,*

$$\lim_{N\to\infty} \max_{\eta\in\mathcal{F}} \mathbb{E}_\eta \Big[\int_0^t \mathbf{1}\{\eta^{\mathcal{F}}_{s\beta_N^+} \in \Delta_*\}\, ds \Big] \;=\; 0 \,.$$

Proof: Since $\beta_N/\beta_N^+ \to 0$, a change of variables in the time integral, similar to the one performed in the proof of Lemma 8.2, and the Markov property show that for every $\eta \in \mathcal{F}$, every $T > 0$ and every N large enough,

$$\mathbb{E}_\eta \Big[\int_0^t \mathbf{1}\{\eta^{\mathcal{F}}_{s\beta_N^+} \in \Delta_*\}\, ds \Big] \;\le\; \frac{2t}{T} \max_{\xi\in\mathcal{F}} \mathbb{E}_\xi \Big[\int_0^T \mathbf{1}\{\eta^{\mathcal{F}}_{s\beta_N} \in \Delta_*\}\, ds \Big] \,.$$

Note that the process on the right hand side is speeded up by β_N instead of β_N^+.

We estimate the expression on the right hand side of the previous formula. We may, of course, restrict the maximum to Δ_*. Let T_1 be the first time the trace process $\eta^{\mathcal{F}}_t$ hits $\mathcal{G}$ and let T_2 be the time it takes for the process to return to Δ_* after T_1:

$$T_1 \;=\; H_{\mathcal{G}} \,, \quad T_2 \;=\; \inf \big\{ s > 0 : \eta^{\mathcal{F}}_{T_1+s} \in \Delta_* \big\} \,.$$

Fix $\eta \in \Delta_*$ and note that

$$\begin{aligned} &\mathbb{E}_\eta \Big[\frac{1}{T} \int_0^T \mathbf{1}\{\eta^{\mathcal{F}}_{s\beta_N} \in \Delta_*\}\, ds \Big] \\ &\qquad \le\; \mathbb{P}^{\mathcal{F}}_\eta \big[T_1 > t_0\beta_N \big] \;+\; \mathbb{P}^{\mathcal{F}}_\eta \big[T_2 \le T\beta_N \big] \;+\; \frac{t_0}{T} \end{aligned}$$

for all $t_0 > 0$. By Assertion 8.E, the first term on the right hand side vanishes as $N \uparrow \infty$ and then $t_0 \uparrow \infty$. On the other hand, by the strong Markov property, the second term is bounded by $\max_{\xi\in\mathcal{G}} \mathbb{P}^{\mathcal{F}}_\xi [H_{\Delta_*} \le T\beta_N]$. Since, by Proposition 6.1, the process $\psi_{\mathcal{F}}(\eta^{\mathcal{F}}_{t\beta_N})$ converges in the Skorohod topology to the Markov chain $X_{\mathcal{F}}(t)$,

$$\limsup_{N\to\infty} \max_{\xi\in\mathcal{G}} \mathbb{P}^{\mathcal{F}}_\xi [H_{\Delta_*} \le T\beta_N] \;\le\; \max_{1\le a\le \mathfrak{q}} \max_{x\in G_a} \mathbf{P}_x [H_{G_{\mathfrak{q}+1}} \le T] \,,$$

where, as in the proof of the previous assertion, $\mathbf{P}_x$ represents the distribution of the chain $X_{\mathcal{F}}(t)$ starting from x. Since the sets G_a are recurrent classes for the chain $X_{\mathcal{F}}(t)$, $r_{\mathcal{F}}(x, y) = 0$ for all $x \in \cup_{1\le a\le \mathfrak{q}} G_a$, $y \in G_{\mathfrak{q}+1}$. Therefore, the previous probability is equal to 0 for all $T > 0$, which completes the proof of the lemma. □

References

L. Avena and A. Gaudillière. On some random forests with determinantal roots. *ArXiv Mathematics e-prints* (2013). arXiv: 1310.1723.

J. Beltrán and C. Landim. Tunneling and metastability of continuous time Markov chains. *J. Stat. Phys.* **140** (6), 1065–1114 (2010). MR2684500.

J. Beltrán and C. Landim. Metastability of reversible finite state Markov processes. *Stochastic Process. Appl.* **121** (8), 1633–1677 (2011). MR2811018.

J. Beltrán and C. Landim. Tunneling and metastability of continuous time Markov chains II, the nonreversible case. *J. Stat. Phys.* **149** (4), 598–618 (2012). MR2998592.

J. Beltrán and C. Landim. A martingale approach to metastability. *Probab. Theory Related Fields* **161** (1-2), 267–307 (2015). MR3304753.

O. Benois, C. Landim and M. Mourragui. Hitting times of rare events in Markov chains. *J. Stat. Phys.* **153** (6), 967–990 (2013). MR3131505.

A. Bianchi and A. Gaudillière. Metastable states, quasi-stationary distributions and soft measures. *Stochastic Process. Appl.* **126** (6), 1622–1680 (2016). MR3483732.

M. Cameron and E. Vanden-Eijnden. Flows in complex networks: theory, algorithms, and application to Lennard-Jones cluster rearrangement. *J. Stat. Phys.* **156** (3), 427–454 (2014). MR3217531.

P. Chleboun and S. Grosskinsky. A dynamical transition and metastability in a size-dependent zero-range process. *J. Phys. A* **48** (5), 055001, 12 (2015). MR3300272.

E. N. M. Cirillo, F. R. Nardi and J. Sohier. Metastability for general dynamics with rare transitions: escape time and critical configurations. *J. Stat. Phys.* **161** (2), 365–403 (2015). MR3401022.

Weinan E and E. Vanden-Eijnden. Towards a theory of transition paths. *J. Stat. Phys.* **123** (3), 503–523 (2006). MR2252154.

R. Fernandez, F. Manzo, F. R. Nardi and E. Scoppola. Asymptotically exponential hitting times and metastability: a pathwise approach without reversibility. *Electron. J. Probab.* **20**, Paper No. 122, 37 (2015). MR3425542.

R. Fernandez, F. Manzo, F. R. Nardi, E. Scoppola and J. Sohier. Conditioned, quasi-stationary, restricted measures and escape from metastable states. *Ann. Appl. Probab.* **26** (2), 760–793 (2016). MR3476624.

M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, second edition (1998). ISBN 0-387-98362-7. MR1652127.

D. Gabrielli and C. Valente. Which random walks are cyclic? *ALEA Lat. Am. J. Probab. Math. Stat.* **9**, 231–267 (2012). MR2923192.

A. Gaudillière and C. Landim. A Dirichlet principle for non reversible Markov chains and some recurrence theorems. *Probab. Theory Related Fields* **158** (1-2), 55–89 (2014). MR3152780.

T. Komorowski, C. Landim and S. Olla. *Fluctuations in Markov processes. Time symmetry and martingale approximation*, volume 345 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg (2012). ISBN 978-3-642-29879-0. MR2952852.

C. Landim. Metastability for a non-reversible dynamics: the evolution of the condensate in totally asymmetric zero range processes. *Comm. Math. Phys.* **330** (1), 1–32 (2014). MR3215575.

C. Landim. A topology for limits of Markov chains. *Stochastic Process. Appl.* **125** (3), 1058–1088 (2015). MR3303968.

J. Lu and E. Vanden-Eijnden. Exact dynamical coarse-graining without time-scale separation. *J. Chem. Phys.* **141** (4), 044109 (2014). DOI: 10.1063/1.4890367.

C. Maes and W. O'Kelly de Galway. A low temperature analysis of the boundary driven Kawasaki process. *J. Stat. Phys.* **153** (6), 991–1007 (2013). MR3131506.

F. Manzo, F. R. Nardi, E. Olivieri and E. Scoppola. On the essential features of metastability: tunnelling time and critical configurations. *J. Statist. Phys.* **115** (1-2), 591–642 (2004). MR2070109.

P. Metzner, C. Schütte and E. Vanden-Eijnden. Transition Path Theory for Markov Jump Processes. *SIAM Multiscale Model. Simul.* **7** (3), 1192–1219 (2009). DOI: 10.1137/070699500.

R. Misturini. Evolution of the ABC model among the segregated configurations in the zero-temperature limit. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** (2), 669–702 (2016). MR3498005.

E. Olivieri and E. Scoppola. Markov chains with exponentially small transition probabilities: first exit problem from a general domain. I. The reversible case. *J. Statist. Phys.* **79** (3-4), 613–647 (1995). MR1327899.

E. Olivieri and E. Scoppola. Markov chains with exponentially small transition probabilities: first exit problem from a general domain. II. The general case. *J. Statist. Phys.* **84** (5-6), 987–1041 (1996). MR1412076.

E. Olivieri and M. E. Vares. *Large deviations and metastability*, volume 100 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge (2005). ISBN 0-521-59163-5. MR2123364.

E. Scoppola. Renormalization group for Markov chains and application to metastability. *J. Statist. Phys.* **73** (1-2), 83–121 (1993). MR1247859.

A. Singer, R. Erban, I. G. Kevrekidis and R. R. Coifman. Detecting intrinsic slow variables in stochastic dynamical systems by anisotropic diffusion maps. *Natl. Acad. Sci. USA* **106** (38), 16090–16095 (2009). DOI: 10.1073/pnas.0905547106.