Metastable Markov chains

Claudio Landim

IMPA, Estrada Dona Castorina 110, CEP 22460, Rio de Janeiro, Brasil
CNRS UMR 6085, Université de Rouen, France
e-mail: landim@impa.br

Abstract: We review recent results on the metastable behavior of continuous-time Markov chains derived through the characterization of Markov chains as unique solutions of martingale problems.

Contents

We present in this review recent developments in the theory of metastable Markov chains. The goal of the theory consists in describing the evolution of a Markov chain by a simpler dynamics, typically one whose state-space is much smaller than the original one, preserving the "macroscopic" features of the original process.

To illustrate the problem, we present in the next section an example which motivates the definitions of metastability introduced in Section 2. We then develop three general methods, based on the characterization of Markov chains as solutions of a martingale problems, to derive the metastable behavior of these dynamics.

There are two recent and compulsory monographs on this subject. The first one, by Olivieri and Vares [111], addresses the problem from the perspective of the large deviations theory, and the second one, by Bovier and Den Hollander [31], uses potential theoretic tools. We do not recall these approaches here and refer the reader to the books. The reader will also find there physical motivations, an historical account and an exhaustive list of references, three aspects which are overlooked here. We tried, though, to include in the references the articles published after 2015.

Throughout the article, all new notation and concepts are introduced in blue. We believe this will help the reader who may want to skip some introductory parts. We present in Section A and B all results on Markov chains and potential theory used in the article. Comments on the method presented in this review are left to the end of Subsection 2.3.

1. A random walk in a graph

We present in this section an example of a Markov chain to motivate three different definitions of metastability. Denote by E_N, $N \geq 1$, the set shown in Figure 1. In this picture, each large square represents a d-dimensional discrete cube of length N, $\Lambda_N = \{1, \ldots, N\}^d$, $d \geq 2$. Each pair of neighboring cubes has one and only one common point. In particular, E_N has $4(N^d - 1)$ elements. Elements of E_N are represented by the Greek letters η, ξ, ζ, and are called points or configurations.
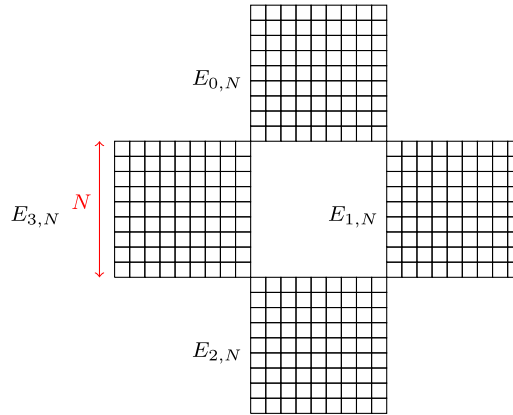
FIG 1. *The set E_N.*

Let $E_{j,N}$, $0 \leq j \leq 3$, be copies of Λ_N. The set E_N is formed by the union of the sets $E_{j,N}$ in which some corner points have been identified. We denote by $E_{0,N}$ the north cube and proceed labeling the sets in the clockwise order so that $E_{3,N}$ represents the west cube.

Denote by $\eta_N(t)$ the continuous-time, E_N-valued, Markov chain which waits a mean-one exponential time at each configuration and then jumps uniformly to one of the neighbor points. This Markov chain is clearly irreducible. Denote by $\deg(\eta)$, $\eta \in E_N$, the degree of the configuration η, that is the number of neighbors. The measure π_N, defined by $\pi_N(\eta) = Z_N^{-1} \deg(\eta)$, where Z_N is the normalizing constant which turns π_N a probability measure, satisfies the detailed balance conditions, and is therefore the unique stationary state.

The purpose of this section is to provide a *synthetic description* of the Markov chain $\eta_N(t)$. In this example, the reduced model is evident. Denote by $\Upsilon_N :$ $E_N \to \{0, 1, 2, 3\}$ the projection which sends a configuration in $E_{j,N}$ to j:

$$\Upsilon_N(\eta) \ = \ \sum_{k=0}^{3} k \, \chi_{E_{k,N}}(\eta) \, ,$$

where χ_A stands for the indicator function of the set A. The value of Υ_N at the intersections of the cubes is not important and can be set arbitrarily.

The derivation of the asymptotic evolution of the *coarse-grained model*

$$Y_N(t) \; = \; \Upsilon_N(\eta_N(t))$$

is based on properties of random walks evolving on discrete cubes. Denote by $z_N(t)$ the symmetric, continuous-time random walk on Λ_N [the process $\eta_N(t)$ restricted to Λ_N], and by π_{Λ_N} its stationary state, the probability measure which gives weights proportional to the degree of the vertices. It is well known, cf. [96, Proposition 10.13], that the mixing time of $z_N(t)$ is of order N^2 and that the time needed to hit a point at distance N is of order $\alpha_N = N^2 \log N$ in dimension 2, and $\alpha_N = N^d$ in dimension $d \geq 3$.

Assume that the chain starts at the center of the cube $E_{j,N}$. Denote by B the set of points which belong to more than one cube, called hereafter the *intersection points*, and by H_B^N the hitting time of B:

$$H_B^N \; = \; \inf\{t \geq 0 : \eta_N(t) \in B\} \,.$$

Since the mixing time is of order N^2 and the hitting time H_B^N is of a much larger order, the chain *equilibrates*, or *thermalizes*, before reaching one of the corners of $E_{j,N}$. This mean that the distribution of the chain approaches π_{Λ_N} before attaining B. In particular, $\eta_N(t)$ looses track of its starting point before hitting one of the corners, and it reaches one of the two intersection points with a probability close to $1/2$.

After thermalizing inside the cube $E_{j,N}$, the random walk $\eta_N(t)$ wanders around $E_{j,N}$ for a length of time of order α_N, and then attains a point in the intersection of $E_{j,N}$ with $E_{j\pm1,N}$, where summation is performed modulo 4. Denote this point by ξ, and assume, to fix ideas, that it belongs to $E_{j,N} \cap E_{j+1,N}$.

Fix a sequence $(\ell_N : N \geq 1)$ such that $\ell_N \to \infty$, $\ell_N/N \to 0$. The precise choice of ℓ_N is not important. Denote by V_N the set of points in E_N which are at an Euclidean distance ℓ_N or less from ξ. After hitting ξ, the random walk performs some short excursions from ξ to ξ which remain in V_N. Some of these excursions are contained in the set $E_{j,N}$ and some in $E_{j+1,N}$.

It takes a time of order ℓ_N^2 for $\eta_N(t)$ to escape from V_N, that is, to reach a point in V_N^c, the complement of V_N. Note that ℓ_N^2 is much smaller than α_N and so the escape time from V_N is negligible in this time-scale.

Starting from a point at the external boundary of V_N, it takes a time of order $N^2 \log \ell_N$ in dimension 2 and N^d in dimension $d \geq 3$ to hit again the set B. Since this time is much longer than the mixing time, once in V_N^c, before hitting the set B again, the process equilibrates inside the cube. Thus, we are back to the initial situation, and we can iterate the previous argument to provide a complete description of the evolution of the random walk $\eta_N(t)$ among the cubes.

According to the previous analysis, the evolution of the random walk can be described as follows. Starting from a point not too close from the corners, the random walk equilibrates in the cube from where it starts before it reaches one of the intersection points. Since it has equilibrated, it reaches one of the two boundary points with equal probability. Then, after some short excursion close

to the intersection point, it escapes from the corner to one of the neighboring cubes, with equal probability due to the symmetry of the set E_N. In particular, with probability $1/2$ the random walk returns to the cube from which it came when it hit the intersection point. The escape time being much shorter than the equilibration time, the small excursions around the intersection can be neglected in the asymptotic regime. After escaping, the process equilibrates in the cube where it is and we may iterate the description of the evolution.

Loss of memory being the essence of Markovian evolution, in the time-scale α_N, the coarse-grained, speeded-up process

$$\mathbf{Y}_N(t) := Y_N(t\alpha_N) = \Upsilon_N(\eta_N(t\alpha_N))$$

should evolve as a $S := \{0,1,2,3\}$-valued, continuous-time Markov chain $\mathbf{Y}(t)$ with holding rates equal to some $\lambda > 0$ and jump probabilities given by $p(j, j \pm 1) = 1/2$.

In which sense can $\mathbf{Y}_N(t)$ converge to a Markov chain? Figure 2 presents a typical realization of the process $\mathbf{Y}_N(t)$. The process remains a time interval of order α_N at a point $x \in S$ until $\eta_N(t)$ reaches an intersection point. At this time, $\eta_N(t)$ performs very short excursions [in the time scale α_N] in both neighboring squares. These short excursions are represented in Figure 2 by the bold rectangles to indicate a large number of oscillations in a very short time interval. After many short excursions the random walk escapes from the boundary and remains in one of the neighboring cubes for a new time interval of order α_N.
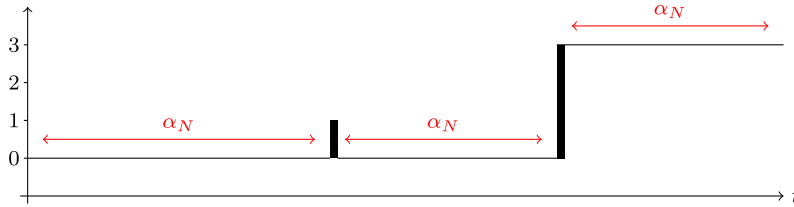
FIG 2. *A typical trajectory of the process $X_N(t)$. The red arrows indicate the length of the time intervals which are of order α_N.*

These fluctuations in very short time intervals, represented by the black rectangles in Figure 2, rule out the possibility that $\mathbf{Y}_N(t)$ converges in any of the Skorohod topologies. Thus, either we content ourselves with the convergence of the finite-dimensional distributions or we need to adjust the trajectories of $\mathbf{Y}_N(t)$ by removing these short excursions.

The first step consists in introducing a set $\Delta_N \subset E_N$ to separate the squares $E_{j,N}$. This procedure is illustrated in Figure 3, where $\mathcal{E}_N^j$ represents $E_{j,N} \setminus \Delta_N$. The set Δ_N is not unique. We only require that it is small enough for the fraction of time spent in Δ_N to be negligible, but large enough for the process, starting from a point outside of Δ_N, to equilibrate before it hits an intersection point.

In the example of this section, the set $\mathcal{E}_N^k$ can be the points of $E_{k,N}$ which are at distance at least ℓ_N from the intersection points, or, as in Figure 3, the

set of points at distance greater than ℓ_N from the faces of the cubes. Here, as above, ℓ_N is a sequence such that $\ell_N \to \infty$, $\ell_N/N \to 0$.
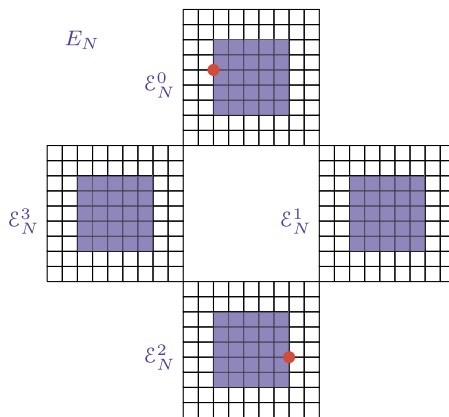
F$_{IG}$ 3. *The sets $\mathcal{E}_N^k$ are indicated in blue. The two red dots represent points in $\mathcal{E}_N^0$ and $\mathcal{E}_N^2$. The trace process $\eta^{\mathcal{E}}(t)$ may jump from one to the other. It has therefore long jumps, in contrast with the original random walks which only jumps to nearest neighbors. The picture is misleading as the annulus around each blue square is much smaller than the square.*

In the next section, we propose two different types of amendments of the trajectories of $\eta_N(t)$ to achieve convergence in the Skorohod topology of the coarse-grained model.

Before we turn to that, consider the example shown in Figure 4. Assume that each line has N points, counting the common intersection point. Consider a random walk evolving on this graph. The process waits a mean-one exponential time at the end of which it jumps to one of its neighbors with equal probability. Since one-dimensional random walks on a set of N points equilibrate in a time of order N^2, and since it hits a point a distance N in the same time-scale, there is no *separation of scales* and the argument presented above to claim the possibility of a synthetic description of the dynamics does not apply.
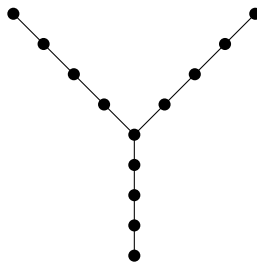
F$_{IG}$ 4. *A random walk on a graph which does not have a synthetic description as a 3-state Markov chain.*

The article is organized as follows. In Section 2, we propose three different definitions of metastability. We present, in Sections 4–7, a general scheme to derive the metastable behavior of a Markov chain in the sense of Definition 2.2 for dynamics which "visit points". This approach is based on the characterization of Markov chains as solutions of martingale problems, examined in Section 3. In the following two sections, an alternative approach is proposed for dynamics in which the entropy plays a role in the metastable behavior. In Section 10, we discuss tightness. In Section 11, we show that conditions (T1), (T2) entail the metastability in the sense of the last passage, and, in Section 12, we prove that these conditions together with property (12.1) lead to the convergence of the finite-dimensional distributions. In Section A and B we recall some general results on Markov chains and potential theory used in the article. In the last section, we list some dynamics which fall within the scope of the theory.

2. Metastability as model reduction

The phenomenon described in the previous section, in which a process remains a long time in a set in which it equilibrates before it attains, in a very short transition, another set where the same behavior is observed, is shared by many different types of dynamics (cf. Section 13 for many examples).

For this reason, we present in a general framework the adjustments needed in the trajectory of the coarse-grained model to yield convergence in the Skorohod topology. Let $(E_N : N \geq 1)$ be a sequence of finite state spaces. Elements of E_N are represented by the Greek letters η, ξ, ζ. Denote by $\eta_N(t)$ a continuous-time, E_N-valued, irreducible Markov chain. Its generator is represented by $\mathcal{L}_N$ and its unique stationary state by π_N. Therefore, for every function $f : E_N \to \mathbb{R}$,

$$(\mathcal{L}_N f)(\eta) \;=\; \sum_{\xi \in E_N} R_N(\eta, \xi) \left[f(\xi) - f(\eta) \right] ,$$

where $R_N(\eta, \xi)$ stands for the jump rates.

For a nonempty subset A of E_N, let H_A, resp. H_A^+, stands for the hitting time of the set A, resp. the return time to A,

$$H_A \;:=\; \inf\{t \geq 0 : \eta_N(t) \in A\} , \quad H_A^+ \;=\; \inf\{t \geq \tau_1 : \eta_N(t) \in A\} . \qquad (2.1)$$

in this formula, τ_1 represents the time of the first jump of $\eta_N(t)$, $\tau_1 = \inf\{t \geq 0 : \eta_N(t) \neq \eta_N(0)\}$.

Assume that E_N contains $\mathfrak{n} > 1$ disjoint sets $\mathcal{E}_N^1, \ldots, \mathcal{E}_N^{\mathfrak{n}}$, called *valleys*, separated by a set Δ_N, so that $\mathcal{E}_N^1, \ldots, \mathcal{E}_N^{\mathfrak{n}}, \Delta_N$ forms a partition of E_N. Let $S := \{1, \ldots, \mathfrak{n}\}$, and denote by $\Phi_N : E_N \to S \cup \{\mathfrak{d}\}$ the projection which sends a configuration in $\mathcal{E}_N^j$, Δ_N to j, $\mathfrak{d}$, respectively:

$$\Phi_N(\eta) \;:=\; \sum_{k=1}^{\mathfrak{n}} k \, \chi_{\mathcal{E}_N^k}(\eta) \;+\; \mathfrak{d} \, \chi_{\Delta_N}(\eta) .$$

Let $X_N(t)$ be the $(S \cup \{\mathfrak{d}\})$-valued process given by

$$X_N(t) \;:=\; \Phi_N\big(\eta_N(t)\big) . \qquad (2.2)$$

In the example of the previous section, the trajectory of $X_N(t) = \Phi_N(\eta_N(t))$ resembles the one presented in Figure 2 with additional spikes due to very short excursions [in the time scale α_N] out of $\mathcal{E}_N^k$ which occur far from the intersection points.

2.1. Last passage

The first adjustment of the trajectories which enables convergence in the Skorohod topology consists in removing the fast fluctuations by recording the last set $\mathcal{E}_N^k$ visited by $\eta_N(t)$. For $t > 0$, denote by $\eta_N(t-)$ the left limit of η_N at t:

$$\eta_N(t-) \;=\; \lim_{s \to t,\, s<t} \eta_N(s)\,.$$

Let $X_N^V(t)$ be given by

$$X_N^V(t) \;:=\; \Phi_N\big(\eta_N(\mathfrak{v}_N(t))\big)\,. \tag{2.3}$$

where

$$\mathfrak{v}_N(t) \;=\; \begin{cases} t & \text{if } \eta_N(t) \in \mathcal{E}_N\,, \\ \mathfrak{w}_N(t)- & \text{otherwise}\,, \end{cases}$$

and $\mathfrak{w}_N(t)$ represents the last time before t the process was in one of the valleys $\mathcal{E}_N^k$:

$$\mathfrak{w}_N(t) \;:=\; \sup\{s \le t : \eta_N(s) \in \mathcal{E}_N\} \quad \text{and} \quad \mathcal{E}_N \;:=\; \bigcup_{k=1}^{\mathfrak{n}} \mathcal{E}_N^k\,.$$

If the set on the right-hand side is empty, we set $\mathfrak{w}_N(t) = 0$. This remark is not important as we will always start the process from a configuration in $\mathcal{E}_N$. Note that $X_N^V(t) \in S$ because $\eta_N(\mathfrak{v}_N(t)) \in \mathcal{E}_N$ for all $t \ge 0$ whenever $\eta_N(0) \in \mathcal{E}_N$.

Here is an alternative way, may be simpler, to define the process $X_N^V(t)$. Define inductively the sequence of stopping times $(T_j : j \ge 0)$ by $T_0 = 0$,

$$T_{j+1} \;=\; \inf\big\{\, t \ge T_j : \Phi(\eta_N(t)) \in S \setminus \{\,\Phi(\eta_N(T_j))\,\}\,\big\}\,, \quad j \ge 0\,.$$

Hence, T_{j+1} is the first time after T_j at which $\Phi(\eta_N(t))$ takes a value different from $\mathfrak{d}$ and $\Phi(\eta_N(T_j))$. In other words, T_{j+1} is the first time after T_j at which the process $\eta_N(t)$ visits a new valley. Then, by definition of $X_N^V(t)$, on the interval $[T_j, T_{j+1})$, $X_N^V(t) = \Phi_N\big(\eta_N(T_j)\big)$:

$$X_N^V(t) \;=\; \sum_{j \ge 0} \Phi_N\big(\eta_N(T_j)\big)\,\chi_{[T_j, T_{j+1})}(t)\,.$$

Here, recall, χ_A stands for the indicator function of the set A.

The time change $\mathfrak{v}_N(t)$ removes the rapid oscillations from the trajectory. Indeed, in the example of the previous section assume that the process starts from a configuration in $\mathcal{E}_N^k$, and denote by $\mathcal{E}_N^j$ the next valley visited. Recall that $H_{\mathcal{E}_N^j}$ represents the hitting time of this valley. In the time interval $[0, H_{\mathcal{E}_N^j})$,

during the rapid excursions of the random walk $\eta_N(t)$ in Δ_N, $X_N^V(t)$ remains equal to k. In particular, the fast fluctuations in the time interval $[0, H_{\mathcal{E}_N^j}]$ are washed out. We may iterate the argument starting from time $H_{\mathcal{E}_N^j}$ to extend this property to the full trajectory.

Since $H_{\mathcal{E}_N^j}$ is of order α_N, the trajectory of $X_N^V(t)$ is formed by a sequence of time intervals of this magnitude in which the process remains constant. The objections raised above for the convergence in the Skorohod topology are thus overturned, and we may expect, due to the loss of memory which emerges from the equilibration, that in the time scale α_N, $X_N^V(t)$ converges to a S-valued Markov chain in the Skorohod topology.

Definition 2.1 (Metastability according to LP). *The Markov chain $\eta_N(t)$ is said to be metastable, in the sense of last passage, in the time-scale θ_N if there exists a partition $\{\mathcal{E}_N^1, \ldots, \mathcal{E}_N^{\mathfrak{n}}, \Delta_N\}$ of the state space E_N and a S-valued, continuous-time Markov chain $\boldsymbol{X}(t)$ such that*

(LP1) *For any $k \in S = \{1, \ldots, \mathfrak{n}\}$ and any sequence $(\eta_N : N \geq 1)$ such that $\eta_N \in \mathcal{E}_N^k$, starting from η_N, $\boldsymbol{X}_N^V(t) = X_N^V(t\,\theta_N)$ converges in the Skorohod topology to $\boldsymbol{X}(t)$.*

(LP2) *The time spent in Δ_N is negligible: For all $t > 0$*

$$\lim_{N \to \infty} \max_{\eta \in \mathcal{E}_N} \mathbb{E}_\eta^N \Big[\int_0^t \chi_{\Delta_N}\big(\eta_N(s\theta_N)\big)\, ds \Big] \;=\; 0\,.$$

The sets $\mathcal{E}_N^j$ are called valleys and the process $\boldsymbol{X}(t)$ the reduced model.

The main difficulty in proving such a result lies in the fact that the process $\eta_N(\mathfrak{v}_N(t))$ is not markovian. For this reason we propose an alternative modification of the trajectory which keeps this property. This method requires the definition of the trace of a process, which we present below in the context of continuous-time Markov chains taking values in a finite state space.

2.2. Trace process

Let E be a finite set and let $\eta(t)$ be an irreducible, continuous-time, E-valued Markov chain. Denote by $R(\eta, \xi)$, $\eta \neq \xi \in E$, the jump rates of this chain, by $\lambda(\eta) = \sum_{\xi \in E} R(\eta, \xi)$ the holding rates, and by π the unique stationary probability measure.

Denote by $D([0, \infty), E)$ the space of right-continuous trajectories $\omega : [0, \infty) \to E$ which have left limits endowed with the Skorohod topology [26]. This notation will be used below, without further comments, replacing E by another metric space. Let $\mathbb{P}_\eta$, $\eta \in E$, be the probability measures on $D([0, \infty), E)$ induced by the Markov chain $\eta(t)$ starting from η. Expectation with respect to $\mathbb{P}_\eta$ is represented by $\mathbb{E}_\eta$.

Fix a non-empty, proper subset F of E and denote by $T_F(t)$, $t \geq 0$, the total

time the process $\eta(t)$ spends in F on the time-interval $[0, t]$:

$$T_F(t) := \int_0^t \chi_F(\eta(s)) \, ds \,,$$

where, we recall, χ_F represents the indicator function of the set F. Denote by $S_F(t)$ the generalized inverse of the additive functional $T_F(t)$:

$$S_F(t) := \sup\{s \geq 0 : T_F(s) \leq t\} \,. \tag{2.4}$$

The irreducibility guarantees that for all $t > 0$, $S_F(t)$ is finite almost surely.

The process T_F is continuous. It is either constant, when the chain visits configurations which do not belong to F, or it increases linearly. Figure 5 illustrates this behavior. Denote by $\eta_F(t)$ the *trace* of the chain $\eta(t)$ on the set F, defined by $\eta_F(t) := \eta(S_F(t))$. Taking the trace of the process corresponds to changing the axis of time in Figure 5. When the process hits F^c, time is frozen until $\eta(t)$ reaches F again, at which time the clock is restarted. In particular, $\eta_F(t)$ takes values in the set F.
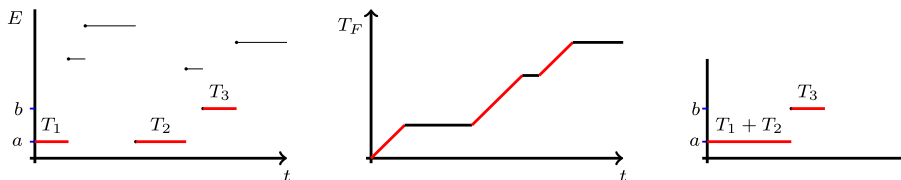
FIG 5. *An example of the transformation which maps the chain $\eta(t)$ into its trace on the set $\{a, b\}$. The first graph shows the trajectory of $\eta(t)$, the second one the function $T_F(t)$ for $F = \{a, b\}$, and the third one the trajectory $\eta_F(t) = \eta(S_F(t))$. Note that $S_F(t)$ is obtained from $T_F(t)$ by inverting the roles of the x and y axes.*

It can be proven [15, Section 6] that $\eta_F(t)$ is an irreducible, continuous-time, F-valued Markov chain. The jump rates of the chain $\eta_F(t)$, denoted by $R_F(\eta, \xi)$, are given by

$$R_F(\eta, \xi) := \lambda(\eta) \, \mathbb{P}_\eta[H_F^+ = H_\xi] \,, \quad \eta, \xi \in F \,, \quad \eta \neq \xi \,, \tag{2.5}$$

where the hitting time H_A and the return time H_A^+ have been introduced in (2.1).

The unique stationary probability measure of the trace chain, denoted by $\pi_F(\eta)$, is the measure π conditioned to F:

$$\pi_F(\eta) = \frac{\pi(\eta)}{\pi(F)} \,, \quad \eta \in F \,. \tag{2.6}$$

Moreover, π_F is reversible if so is π [15].

2.3. Metastability

We return to the chain $\eta_N(t)$ introduced at the beginning of this section. Denote by $\mathbb{P}_\eta^N$, $\eta \in E_N$, the probability measures on $D([0,\infty), E_N)$ induced by the Markov chain $\eta_N(t)$ starting from η. Expectation with respect to $\mathbb{P}_\eta^N$ is represented by $\mathbb{E}_\eta^N$.

Denote by $\eta^{\mathcal{E}_N}(t)$ the trace of the process $\eta_N(t)$ on the set $\mathcal{E}_N$. As explained in Figure 5, by taking the trace of $\eta_N(t)$ on $\mathcal{E}_N$ we first remove from the trajectory the time-intervals corresponding to the excursions in Δ_N (the intervals in black in the leftmost picture of Figure 5), and then, we push back the trajectory, as in the rightmost picture of this figure. This procedure removes rapid fluctuations from the trajectory providing an alternative definition of metastability.

Let $\Psi_N \colon \mathcal{E}_N \to S$ the projection which sends a configuration in $\mathcal{E}_N^j$ to j:

$$\Psi_N(\eta) \;=\; \sum_{k=1}^{\mathfrak{n}} k \, \chi_{\mathcal{E}_N^k}(\eta) \,.$$

In contrast with Φ_N, Ψ_N is defined only on $\mathcal{E}_N$. Let $X_N^T(t)$ be the process given by

$$X_N^T(t) \;:=\; \Psi_N\big(\eta^{\mathcal{E}_N}(t)\big) \,. \tag{2.7}$$

Note that $X_N^T(t)$ is not a Markov chain, but just a hidden Markov chain. It corresponds to the trace on S of the process $X_N(t)$ introduced in (2.2).

Definition 2.2 (Metastability). *The Markov chain $\eta_N(t)$ is said to be metastable in the time-scale θ_N if there exists a partition $\{\mathcal{E}_N^1, \ldots, \mathcal{E}_N^{\mathfrak{n}}, \Delta_N\}$ of the state space E_N and a S-valued, continuous-time Markov chain $\boldsymbol{X}(t)$ such that*

(T1) *For any $k \in S = \{1, \ldots, \mathfrak{n}\}$ and any sequence $(\eta_N : N \geq 1)$ such that $\eta_N \in \mathcal{E}_N^k$, starting from η_N, the process $\boldsymbol{X}_N^T(t) := X_N^T(t\theta_N) = \Psi_N(\eta^{\mathcal{E}_N}(t\theta_N))$ converges in the Skorohod topology to $\boldsymbol{X}(t)$;*

(T2) *The time spent in Δ_N is negligible: For all $t > 0$*

$$\lim_{N\to\infty} \max_{\eta\in\mathcal{E}_N} \mathbb{E}_\eta^N\Big[\int_0^t \chi_{\Delta_N}\big(\eta_N(s\theta_N)\big)\, ds\Big] \;=\; 0 \,.$$

The first condition asserts that in the time scale θ_N the trace on S of the process $X_N(t)$ converges to a Markov chain, while the second one states that in this time scale the amount of time the process $X_N(t)$ spends outside S is negligible, uniformly over initial configurations in $\mathcal{E}_N$. In particular, condition (T2) can be stated as

$$\lim_{N\to\infty} \max_{\eta\in\mathcal{E}_N} \mathbb{E}_\eta^N\Big[\int_0^t \chi_{\mathfrak{d}}\big(X_N(s\theta_N)\big)\, ds\Big] \;=\; 0 \,.$$

Remark 2.3. *The use of the word "metastability", instead of tunneling, to name the phenomenon described in the previous section, might be inadequate. Metastability has been used to represent the transition from a metastable state to*

a stable one. This corresponds to the case in which the reduced model $\boldsymbol{X}(t)$ takes value in a set with two elements, one being transient and the other absorbing. We allow ourselves this abuse of nomenclature.

Remark 2.4. *The same sequence of Markov chains $(\eta_N(t) : N \geq 1)$ may have more that one metastable description. In a certain time-scale α_N, one may observe transitions between shallow valleys and in a much longer time-scale β_N transitions between deeper valleys.*

Remark 2.5. *There are examples of Markov chains [63, 62, 78, 79, 13] with a countably infinite number of valleys. In these cases, the reduced model $\boldsymbol{X}(t)$ is a continuous-time Markov chain in a countable state-space. In this article, we restrict ourselves to the finite case to avoid technical issues on the martingale problem.*

Remark 2.6. *One of the main features of metastability is the fast transition between valleys. This information is encapsulated in condition (T2) which states that the time spent outside the valleys is negligible. In particular, the transition time between two valleys is negligible in the metastable time-scale.*

Remark 2.7. *All results presented in this review are in asymptotic form, they characterize the limiting behavior of the coarse-grained model. Quantitative estimates at fixed N are important in concrete problems. For example, to describe synthetically a molecular dynamics which can be represented as a Markov chain in a very large, but fixed, state space. The problem consists in finding a reduced model which keeps the main features of the original chain. It might be interesting to adapt the approach presented here to this framework.*

The transition path theory [54, 101, 33, 98] has been designed for this set-up, as well as the intertwining method [11, 8, 9, 10]. See also the results by Bianchi and Gaudillière [24]

Remark 2.8. *In constrast with the pathwise approach [38, 111], no attempt is made here to describe the transition path between two valleys.*

Remark 2.9. *In the example of the previous section, the process $X_N(t)$ remains constant in time-intervals of length of order α_N. In this sense, Ψ_N can be understood as a* slow variable, *since it evolves in a much longer time-scale than the original process, and metastability as the search for slow variables and the description of the evolution of these slow variables.*

Remark 2.10. *In most examples, as the Ising model at low temperature [104, 105], metastability is observed as a result of the presence of an energy barrier which the system has to overpass to reach a new region of the state-space.*

The example of the previous section is of different nature. In this model, there is no energy landscape but a bottleneck which creates a metastable behavior. Here, entropy [the number of configurations] determines the height of the barriers. Say, for example, that three squares are 3-dimensional while the last one is 2-dimensional. In this case, in the time-scale N^3, one observes an evolution among the 3-dimensional cubes and the last square can be included in the set Δ_N as the time spent there is of order $N^2 \log N$.

In other models, as random walks in a potential field, both energy and entropy play a role.

2.4. Finite-dimensional distributions

Definition 2.1 describes the evolution of a modified version of the original process, and Definition 2.2 the one of the trace. To avoid tiny surgeries of the trajectories, we may turn to the convergence of the finite-dimensional distributions, an alternative adopted by Kipnis and Newman in [81] and Sugiura [121, 122].

Definition 2.11 (Metastability according to FDD). *The Markov chain $\eta_N(t)$ is said to be metastable, in the sense of finite-dimensional distributions, in the time-scale θ_N if there exists a partition $\{\mathcal{E}_N^1, \ldots, \mathcal{E}_N^{\mathfrak{n}}, \Delta_N\}$ of the state space E_N and a S-valued, continuous-time Markov chain $\boldsymbol{X}(t)$ such that the finite-dimensional distributions of $\boldsymbol{X}_N(t) := X_N(t\theta_N)$ converge to the ones of $\boldsymbol{X}(t)$.*

Note that while $\boldsymbol{X}_N(t)$ takes value in $S \cup \{\mathfrak{d}\}$, $\boldsymbol{X}(t)$ is S-valued.

3. Martingale problems

The proof of condition (T1) in Definition 2.2 relies on the uniqueness of solutions of martingale problems, the subject of this section. To avoid technical problems, we restrict ourselves to the context continuous-time Markov chains taking values in a finite state-space E. We refer to the classical books [120, 55] for further details.

Recall the notation introduced in Subsection 2.2. Assume that the Markov chain $\eta(t)$ is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\Omega = D([0, \infty), E)$ and $\mathcal{F}$ represents the Borel σ-algebra of $D([0, \infty), E)$. Let $(\mathcal{F}_t^o : t \geq 0)$ be the filtration generated by $\{\eta(s) : 0 \leq s \leq t\}$.

Denote by L the generator of the Markov chain $\eta(t)$: for every function $f : E \to \mathbb{R}$,

$$(Lf)(\eta) = \sum_{\xi \in E} R(\eta, \xi) \left[f(\xi) - f(\eta) \right] . \tag{3.1}$$

It is well known that for every $f : E \to \mathbb{R}$,

$$M^f(t) = f(\eta(t)) - f(\eta(0)) - \int_0^t (Lf)(\eta(s)) \, ds \tag{3.2}$$

is a zero-mean martingale in $(\Omega, (\mathcal{F}_t^o), \mathbb{P})$.

It turns out that the converse is true. Let A be the generator of an E-valued, irreducible, continuous-time Markov chain, and ν a probability measure on E.

Definition 3.1 (The martingale problem (A, ν)). *A probability measures $\boldsymbol{P}$ on $(\Omega, \mathcal{F})$ is a solution of the martingale problem associated to the generator A and the measure ν if for every $f : E \to \mathbb{R}$ the process M^f given by (3.2) [with L replaced by A] is a martingale in $(\Omega, (\mathcal{F}_t^o), \boldsymbol{P})$ and $\boldsymbol{P}[\eta(0) = \eta] = \nu(\eta)$ for all $\eta \in E$.*

Next result is a particular case of Theorem 4.4.1 in [55].

Theorem 3.2. *Let A be the generator of an E-valued, irreducible, continuous-time Markov chain. For every probability measure ν in E, there exists a unique solution of the martingale problem associated to the generator A and the measure ν. Moreover, under this solution, the process $\eta(t)$ is the continuous-time Markov chain whose generator is A.*

This result provides a simple strategy to prove condition (T1) of Definition 2.2. Fix $k \in S$, a sequence $\eta_N \in \mathcal{E}_N^k$, and denote by $\mathbf{P}^N$ the probability measure on $D([0,\infty), S)$ induced by the process $\mathbf{X}_N^T(t) = X_N^T(t\theta_N)$ and the measure $\mathbb{P}_{\eta_N}^N$. Prove first that the sequence $\mathbf{P}^N$ is tight. Then, to characterize the limit points, show that they solve a martingale problem (L, δ_k), where L is the generator of a S-valued Markov chain [guessed a priori] and δ_k the probability measure on S concentrated on k. Tightness is postponed to Section 10 and uniqueness is discussed in the next sections.

4. The martingale approach

We carry out in this section the strategy outlined in the previous section to prove the uniqueness of limit points of the sequence $\mathbf{X}_N^T(t)$. It is based on the uniqueness of solutions of martingale problems, presented above, and on the fact that limits of martingales are martingales recalled below.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $(\mathcal{F}_t : t \geq 0)$ a filtration, and $(M_N : N \geq 1)$ a sequence of martingales measurable with respect to the filtration.

Lemma 4.1. *Assume that for each $t \geq 0$, $M_N(t)$ converges in $L^1(\mathbb{P})$ to a random variable $M(t)$. Then, $M(t)$ is a martingale with respect to the filtration $(\mathcal{F}_t : t \geq 0)$.*

Proof. Fix $0 \leq s < t$ and a bounded random variable Y, measurable with respect to $\mathcal{F}_s$. Since M_N is a martingale,

$$\mathbb{E}\big[\, M_N(t)\, Y\,\big] \;=\; \mathbb{E}\big[\, M_N(s)\, Y\,\big]\,.$$

As Y is bounded and $M_N(t)$, $M_N(s)$ converge in $L^1(\mathbb{P})$ to $M(t)$, $M(s)$, respectively. The same identity holds with M_N replaced by M. Moreover, $M(t)$, $M(s)$ belong to $L^1(\mathbb{P})$. Since this identity is in force for all bounded random variable Y, $\mathbb{E}[\, M(t)\,|\,\mathcal{F}_s\,] \;=\; M(s)$, as claimed. $\qquad\square$

Fix $k \in S$, a configuration η_N in $\mathcal{E}_N^k$, and denote by $\mathbf{P}^N$ the probability measure on $D([0,\infty), S)$ induced by the process $\mathbf{X}_N^T(t)$ and the measure $\mathbb{P}_{\eta_N}^N$. The main result of this section asserts that all limit points of the sequence $\mathbf{X}_N^T(t)$ solve a martingale problem (L, δ_k) if we can prove a *local ergodic theorem* and calculate the limit of the *coarse-grained jump function*, properties (P1) and (P2) formulated at the end of this section.

Fix a function $F : S \to \mathbb{R}$. As the trace process is a Markov chain, (3.2) applied to the function $F \circ \Psi_N$ yields that under $\mathbb{P}^N_{\eta_N}$

$$M_N(t) \;=\; F\big(\Psi_N(\eta^{\mathcal{E}_N}_{t\theta_N}) \big) \;-\; F\big(\Psi_N(\eta^{\mathcal{E}_N}_0) \big) \;-\; \int_0^{t\theta_N} \big[\mathcal{L}_{\mathcal{E}_N}(F \circ \Psi_N) \big] (\eta^{\mathcal{E}_N}_s) \, ds \tag{4.1}$$

is a martingale. In this formula, $\mathcal{L}_{\mathcal{E}_N}$ represents the generator of the trace process $\eta^{\mathcal{E}_N}(t)$. Since $\boldsymbol{X}^T_N(t) = \Psi_N(\eta^{\mathcal{E}_N}(t\theta_N))$, changing variables this expression becomes

$$F\big(\boldsymbol{X}^T_N(t) \big) \;-\; F\big(\boldsymbol{X}^T_N(0) \big) \;-\; \int_0^t \theta_N \big[\mathcal{L}_{\mathcal{E}_N}(F \circ \Psi_N) \big] (\eta^{\mathcal{E}_N}(s\theta_N)) \, ds \; .$$

Denote by $R^T_N(\eta, \xi)$ the jump rates of the trace chain $\eta^{\mathcal{E}_N}(t)$. The expression inside of the integral can be written as

$$\theta_N \sum_{\xi \in \mathcal{E}_N} R^T_N(\zeta, \xi) \big\{ (F \circ \Psi_N)(\xi) - (F \circ \Psi_N)(\zeta) \big\} \quad \text{where} \quad \zeta = \eta^{\mathcal{E}_N}(s\theta_N) \; .$$

Writing $\mathcal{E}_N$ as $\cup_\ell \mathcal{E}^\ell_N$, since $\Psi_N(\xi) = \ell$ for $\xi \in \mathcal{E}^\ell_N$, this expression is equal to

$$\theta_N \sum_{\ell \in S} \sum_{\xi \in \mathcal{E}^\ell_N} R^T_N\big(\eta^{\mathcal{E}_N}(s\theta_N), \xi \big) \big\{ F(\ell) - F(\boldsymbol{X}^T_N(s)) \big\}$$
$$= \sum_{\ell \in S} R^{(\ell)}_N\big(\eta^{\mathcal{E}_N}(s\theta_N) \big) \big\{ F(\ell) - F(\boldsymbol{X}^T_N(s)) \big\} \; ,$$

where $R^{(\ell)}_N(\zeta)$ represent the jump rate from the configuration ζ to the set $\mathcal{E}^\ell_N$ for the trace process speeded-up by θ_N:

$$R^{(\ell)}_N(\zeta) \;=\; \theta_N \sum_{\xi \in \mathcal{E}^\ell_N} R^T_N(\zeta, \xi) \; . \tag{4.2}$$

Up to this point, we proved that the martingale $M_N(t)$ is equal to

$$F\big(\boldsymbol{X}^T_N(t) \big) - F\big(\boldsymbol{X}^T_N(0) \big) - \int_0^t \sum_{\ell \in S} R^{(\ell)}_N\big(\eta^{\mathcal{E}_N}(s\theta_N) \big) \big\{ F(\ell) - F(\boldsymbol{X}^T_N(s)) \big\} \, ds \; .$$

If the functions $R^{(\ell)}_N$ were constant over the sets $\mathcal{E}^\ell_N$, $R^{(\ell)}_N(\eta) = r^{(\ell)}_N(\Psi_N(\eta))$ for some $r^{(\ell)}_N : S \to \mathbb{R}_+$, the martingale $M_N(t)$ could be written in terms of the process $\boldsymbol{X}^T_N(s)$:

$$M_N(t) = F(\boldsymbol{X}^T_N(t)) - F(\boldsymbol{X}^T_N(0)) - \int_0^t \sum_{\ell \in S} r^{(\ell)}_N(\boldsymbol{X}^T_N(s)) \big\{ F(\ell) - F(\boldsymbol{X}^T_N(s)) \big\} \, ds.$$

Furthermore, if for all $j \neq \ell \in S$, the sequences $r^{(\ell)}_N(j)$ converged to some $\boldsymbol{r}(j, \ell) \in \mathbb{R}_+$, one could replace in the previous formula $r^{(\ell)}_N(\boldsymbol{X}^T_N(s))$ by $\boldsymbol{r}(\boldsymbol{X}^T_N(s), \ell)$ at the cost of a small error.

Therefore, under the two previous conditions, up to a negligible error,

$$F(\boldsymbol{X}_N^T(t)) \,-\, F(\boldsymbol{X}_N^T(0)) \,-\, \int_0^t \sum_{\ell \in S} \boldsymbol{r}(\boldsymbol{X}_N^T(s), \ell) \left\{ F(\ell) - F(\boldsymbol{X}_N^T(s)) \right\} ds \quad (4.3)$$

is a martingale.

Denote by $\mathbf{P}$ a limit point of the sequence $\mathbf{P}^N$. Let $X(t)$ represent the coordinate process of $D([0, \infty), S)$:

$$X(t, \omega) \,=\, \omega(t) \,, \quad \omega \in D([0, \infty), S) \,, \quad t \geq 0 \,.$$

Assume that $\mathbf{P}[X(t-) = X(t)] = 1$ for all $t > 0$, where $X(t-) = \lim_{s < t, s \to t} X(s)$.

Suppose, without loss of generality, that $\mathbf{P}^N$ converges to $\mathbf{P}$. Let L be the generator of the S-valued Markov chain associated to the jump rates r. As $\mathbf{P}[X(t-) = X(t)] = 1$, the finite-dimensional projections are continuous (cf. equation (13.3) in [26]). Thus, since the expression in (4.3) is uniformly bounded, we may pass to the limit and conclude from Lemma 4.1 that

$$F(X(t)) \,-\, F(X(0)) \,-\, \int_0^t (L\,F)(X(s))\, ds$$

is a martingale under the measure $\mathbf{P}$. Moreover, as $\eta_N \in \mathcal{E}_N^k$, $\mathbf{P}^N[X(0) = k] = 1$ for all N so that $\mathbf{P}[X(0) = k] = 1$. Therefore, $\mathbf{P}$ is a solution of the (L, δ_k) martingale problem. By Theorem 3.2, this property characterizes $\mathbf{P}$, and under this measure the coordinate process is a continuous-time Markov chain whose generator is L.

We summarize the conclusions of the previous analysis in Theorem 4.2 below. We first formulate the main hypotheses.

(P1) (Local ergodicity). The mean rate functions $R_N^{(\ell)}(\eta)$, introduced in (4.2), can be replaced by coarse-grained functions $r_N^{(\ell)}$. More precisely, there exist sequences of functions $r_N^{(\ell)} : \mathcal{E}_N \to \mathbb{R}_+$, $\ell \in S$, which are constant on the sets $\mathcal{E}_N^j$, $j \in S$, and such that for every function $F : S \to \mathbb{R}$, $t > 0$, and sequence $\eta_N \in \mathcal{E}_N$,

$$\lim_{N \to \infty} \mathbb{E}_{\eta_N} \left[\int_0^t F(\boldsymbol{X}_N^T(s)) \left\{ R_N^{(\ell)}\big(\eta^{\mathcal{E}_N}(s\theta_N) \big) - r_N^{(\ell)}(\boldsymbol{X}_N^T(s)) \right\} ds \right] \,=\, 0 \,.$$

(P2) (The coarse-grained jump rates). The sequence of functions $r_N^{(\ell)}$, $\ell \in S$, called the *coarse-grained jump functions*, converge. More precisely, since these functions are constant over the valleys $\mathcal{E}_N^k$, they can be written as

$$r_N^{(\ell)}(\eta) \,=\, \sum_{k \in S} \boldsymbol{r}_N(k, \ell)\, \chi_{\mathcal{E}_N^k}(\eta)$$

for some non-negative real numbers $\boldsymbol{r}_N(k, \ell)$, named the *coarse-grained jump rates*. Note from the formula for the martingale $M_N(t)$ that the

values of $\boldsymbol{r}_N(\ell, \ell)$ are unimportant. We assume that these rates converge: There exist $\boldsymbol{r}(j, \ell) \in [0, \infty)$, such that for all $j \neq \ell \in S$,

$$\lim_{N \to \infty} \boldsymbol{r}_N(j, \ell) = \boldsymbol{r}(j, \ell) . \tag{4.4}$$

Theorem 4.2. *Fix $k \in S$, a sequence $\eta_N \in \mathcal{E}_N^k$, and denote by $\mathbf{P}^N$ the probability measure on $D([0, \infty), S)$ induced by the process $\boldsymbol{X}_N^T(t)$ and the measure $\mathbb{P}_{\eta_N}^N$. Assume that conditions (P1) and (P2) are in force. Then, every limit point $\mathbf{P}$ of the sequence $\mathbf{P}^N$ such that*

$$\mathbf{P}\big[X(t-) = X(t) \big] = 1 \quad \text{for all } t > 0 .$$

solves the (L, δ_k) martingale problem, where L is the generator of the S-valued Markov chain whose jump rates are $\boldsymbol{r}(j, \ell)$.

Note that we do not need to prove property (P1) with an absolute value inside the expectation. This observation simplifies considerably the proof of this replacement.

We present in Sections 5, 6 sufficient conditions, formulated in terms of the stationary state and of capacities between the sets $\mathcal{E}_N^j$, for conditions (P1), (P2) to hold. In Sections 8, 9 we propose alternative proofs of the uniqueness of limit points for the sequence $\mathbf{P}^N$.

5. Local ergodicity

In this section, we provide sufficient conditions, formulated in terms of the stationary state and of capacities, to replace the jump rates $R_N^{(k)}$, introduced in (4.2), by coarse-grained jump functions which are constant on each set $\mathcal{E}_N^k$. We assume that the reader is acquainted with the results on potential theory of Markov chains, recapitulated in Section B

Recall from (2.1) the definition of the hitting time $H_{\mathcal{A}}$ and the return time $H_{\mathcal{A}}^+$ of a subset $\mathcal{A}$ of E_N. For two non-empty, disjoint subsets $\mathcal{A}$, $\mathcal{B}$ of E_N, denote by $\mathrm{cap}_N(\mathcal{A}, \mathcal{B})$ the capacity between $\mathcal{A}$ and $\mathcal{B}$:

$$\mathrm{cap}_N(\mathcal{A}, \mathcal{B}) = \sum_{\eta \in \mathcal{A}} \pi_N(\eta) \, \lambda_N(\eta) \, \mathbb{P}_{\eta}^N \big[H_{\mathcal{B}} < H_{\mathcal{A}}^+ \big] , \tag{5.1}$$

where $\lambda_N(\eta)$ stands for the holding rate at η of the Markov chain $\eta_N(t)$: $\lambda_N(\eta) = \sum_{\xi \in E_N} R_N(\eta, \xi)$.

Recall that

$$R_N^{(k)}(\eta) = \theta_N \sum_{\xi \in \mathcal{E}_N^k} R_N^T(\eta, \xi) ,$$

where $R_N^T(\eta, \xi)$ represents the jump rates of the trace process. Thus, $R_N^{(k)}(\eta)$ is the rate at which the trace process jumps from η to $\mathcal{E}_N^k$ multiplied by θ_N. In view of equation (2.5) for the jump rates of the trace process,

$$R_N^{(k)}(\eta) = \theta_N \, \lambda_N(\eta) \, \mathbb{P}_{\eta} \big[H_{\mathcal{E}_N}^+ = H_{\mathcal{E}_N^k} \big] = \theta_N \, \lambda_N(\eta) \, \mathbb{P}_{\eta} \big[H_{\mathcal{E}_N^k} < H_{\breve{\mathcal{E}}_N^k}^+ \big] , \tag{5.2}$$

where

$$\breve{\mathcal{E}}_N^k := \bigcup_{j \neq k} \mathcal{E}_N^j , \quad k \in S .$$

In particular, $R_N^{(k)}$ vanishes in the interior of the sets $\mathcal{E}_N^j$, where by interior we mean the set of configurations in $\mathcal{E}_N^j$ whose neighbors belong to $\mathcal{E}_N^j$ [the configuration $\xi \in \mathcal{E}_N^j$ such that $\sum_{\zeta \notin \mathcal{E}_N^j} R_N(\xi, \zeta) = 0$]. This means that $R_N^{(k)}$ is a singular function. While it vanishes in the interior of the sets $\mathcal{E}_N^j$, it assumes a large value at the boundary because the right-hand side of (5.2) is multiplied by θ_N.

The goal of this section is to replace the time integral of the singular function $R_N^{(k)}$ by the time integral of a very regular function, one which is constant at each set $\mathcal{E}_N^j$. This replacement is expected to hold whenever the process equilibrates in the valleys $\mathcal{E}_N^j$ before it jumps to a new one.

Let $f_N : \mathcal{E}_N \to \mathbb{R}$ be a sequence of real functions defined on $\mathcal{E}_N$. Fix $t > 0$, and consider the time integral

$$\int_0^t f_N\big(\eta^{\mathcal{E}_N}(s\theta_N)\big)\, ds = \frac{1}{\theta_N} \int_0^{t\theta_N} f_N\big(\eta^{\mathcal{E}_N}(s)\big)\, ds .$$

The time integral can be decomposed according to the sojourns in the sets $\mathcal{E}_N^j$. If the process equilibrates during these visits, by the ergodic theorem, we expect the integral of f_N over these time-intervals to be close to the integral of the mean value of f_N on these sets. Hence, let

$$\widehat{f}_N(\eta) = E_{\pi_{\mathcal{E}}}\big[f_N \,|\, \mathcal{G}_N \big] , \tag{5.3}$$

where $\mathcal{G}_N$ represents the σ-algebra of subsets of $\mathcal{E}_N$ generated by the sets $\mathcal{E}_N^j$, $j \in S$, and $\pi_{\mathcal{E}}$ the stationary state of the trace process $\eta^{\mathcal{E}_N}(t)$ [which, by (2.6), is the stationary state π_N conditioned to $\mathcal{E}_N$].

Clearly,

$$\widehat{f}_N(\eta) = \sum_{j \in S} F_N(j)\, \chi_{\mathcal{E}_N^j}(\eta) , \quad \text{where} \quad F_N(j) = \frac{1}{\pi_N(\mathcal{E}_N^j)} \sum_{\zeta \in \mathcal{E}_N^j} \pi_N(\zeta)\, f_N(\zeta) .$$

The function $\widehat{f}_N$ is the candidate, and one expects that, under certain conditions on the sequence f_N,

$$\frac{1}{\theta_N} \int_0^{t\theta_N} \left\{ f_N\big(\eta^{\mathcal{E}_N}(s)\big) - \widehat{f}_N\big(\eta^{\mathcal{E}_N}(s)\big) \right\} ds$$

vanishes as $N \to \infty$.

Theorem 5.1. *Let f_N, $g_N : \mathcal{E}_N \to \mathbb{R}$ be sequences of functions such that*

(a) For each $j \in S$, there exists a configuration $\xi^{j,N}$ such that

$$\lim_{N \to \infty} \frac{1}{\theta_N} \max_{\eta \in \mathcal{E}_N^j} \frac{\sum_{\zeta \in \mathcal{E}_N^j} |f_N(\zeta)|\, \pi_N(\zeta)}{\mathrm{cap}_N(\eta, \xi^{j,N})} = 0 ,$$

where the supremum is carried over all configurations $\eta \neq \xi^{j,N}$.

(b) *The sequence g_N is uniformly bounded and is constant over each set $\mathcal{E}_N^j$: There exist a finite constant C_0 and a sequence of functions $G_N : S \to \mathbb{R}$ such that*

$$g_N(\eta) \; = \; \sum_{j \in S} G_N(j) \, \chi_{\mathcal{E}_N^j}(\eta) \quad and \quad \max_{\eta \in \mathcal{E}_N} \big| \, g_N(\eta) \, \big| \; \leq \; C_0$$

for all $N \geq 1$.

Then, for all $t > 0$,

$$\lim_{N \to \infty} \max_{\eta \in \mathcal{E}_N} \Big| \mathbb{E}_\eta^N \Big[\int_0^t \big\{ f_N \big(\eta^{\mathcal{E}_N}(s\theta_N) \big) - \widehat{f}_N \big(\eta^{\mathcal{E}_N}(s\theta_N) \big) \big\} g_N \big(\eta^{\mathcal{E}_N}(s\theta_N) \big) \, ds \Big] \Big| = 0.$$

In the reversible case, this result follows from Corollary 6.5 and Proposition 6.10 in [15] and from the hypotheses of the theorem. In the nonreversible case, it follows from Corollary 6.5 in [15] and Proposition A.2 in [19].

Remark 5.2. *The proof of this result takes advantage of the fact that the absolute value is outside of the expectation.*

Remark 5.3. *To turn the martingale $M_N(t)$, introduced in the previous section, into a function of X_N^T, we only need to prove the previous theorem for $f_N = R_N^{(j)}$, $j \in S$. In this special case, by (5.1) and (5.2),*

$$\frac{1}{\theta_N} \sum_{\zeta \in \mathcal{E}_N^j} | \, f_N(\zeta) \, | \, \pi_N(\zeta) \; = \; \mathrm{cap}_N(\mathcal{E}_N^j, \breve{\mathcal{E}}_N^j) \, .$$

In particular, condition (a) of the theorem becomes that for all $j \in S$, there exists $\xi^{j,N} \in \mathcal{E}_N^j$ such that

$$\lim_{N \to \infty} \max_{\eta \in \mathcal{E}_N^j , \, \eta \neq \xi^{j,N}} \frac{\mathrm{cap}_N(\mathcal{E}_N^j, \breve{\mathcal{E}}_N^j)}{\mathrm{cap}_N(\eta, \xi^{j,N})} \; = \; 0 \, . \tag{5.4}$$

Remark 5.4. *The configuration $\xi^{j,N}$ has no special role. By Theorem 2.7 in [15], if condition (5.4) holds for one configuration in $\mathcal{E}_N^j$, it holds for all.*

The coarse-grained jump function, denoted by $r_N^{(j)}$ in (4.4), is given by $\widehat{R}_N^{(j)}$. Thus, by (5.3) and (5.2),

$$r_N^{(j)}(\eta) \; = \; \widehat{R}_N^{(j)}(\eta) \; = \; E_{\pi_{\mathcal{E}}} \big[R_N^{(j)} \, | \, \mathcal{G}_N \big] \; = \; \sum_{k \in S} \boldsymbol{r}_N(k,j) \, \chi_{\mathcal{E}_N^k}(\eta) \, ,$$

where, for $k \neq j$,

$$\begin{aligned}
\boldsymbol{r}_N(k,j) \; &= \; \frac{1}{\pi_N(\mathcal{E}_N^k)} \sum_{\zeta \in \mathcal{E}_N^k} \pi_N(\zeta) \, R_N^{(j)}(\zeta) \\
&= \; \frac{\theta_N}{\pi_N(\mathcal{E}_N^k)} \sum_{\zeta \in \mathcal{E}_N^k} \pi_N(\zeta) \, \lambda_N(\zeta) \, \mathbb{P}_\zeta \big[\, H_{\mathcal{E}_N^j} < H_{\breve{\mathcal{E}}_N^j}^+ \, \big] \, .
\end{aligned} \tag{5.5}$$

Remark 5.5. *Hypothesis (a) of Theorem 5.1 requires the process to visit all configurations of the valley $\mathcal{E}_N^j$ before it reaches a new one. Dynamics which display this behavior are said to "visit points". This class includes condensing zero-range processes [16, 85, 4, 117], random walks in a potential field [91, 92, 93] or models in which the valleys are singletons as the inclusion process [25] or random walks evolving among random traps [63, 62, 78, 79], but it does not contain the example of Section 1. For such dynamics, in which the entropy plays a role in the metastable behavior, a different approach is needed. This is discussed in Sections 8 and 9.*

6. The coarse-grained jump rates

In this section, we investigate the asymptotic behavior of the coarse-grained jump rates $\boldsymbol{r}_N(k,j)$, defined in (5.5). This is condition (P2) of Theorem 4.2.

6.1. Reversible case

In the reversible case, we may express the coarse-grained jump rates $\boldsymbol{r}_N(j,k)$ in terms of capacities. If follows from the explicit formulae (5.1), (5.5) and from an elementary argument taking advantage of the reversibility that

$$
\begin{aligned}
&\pi_N(\mathcal{E}^j)\,\boldsymbol{r}_N(j,k) \\
&= \theta_N \frac{1}{2} \left\{ \mathrm{cap}_N(\mathcal{E}^j, \breve{\mathcal{E}}^j) + \mathrm{cap}_N(\mathcal{E}^k, \breve{\mathcal{E}}^k) - \mathrm{cap}_N\big(\mathcal{E}^j \cup \mathcal{E}^k\,,\, \cup_{\ell \neq j,k}\mathcal{E}^\ell\big) \right\}.
\end{aligned}
$$

Here and below we often write $\mathcal{E}^j$, $\breve{\mathcal{E}}^j$ for $\mathcal{E}_N^j$, $\breve{\mathcal{E}}_N^j$, respectively. Therefore, in the reversible case, one can compute the limit of the coarse-grained jump rates $\boldsymbol{r}_N(j,k)$ if one can calculate the asymptotic behavior of $\pi_N(\mathcal{E}^j)$ and of

$$
\theta_N \, \mathrm{cap}_N\big(\cup_{j \in A} \mathcal{E}^j\,,\, \cup_{k \in B}\mathcal{E}^k \big)
$$

for non-empty subset A, B of S such that $A \cap B = \varnothing$, $A \cup B = S$.

6.2. Nonreversible case

Summing over $j \neq k$ in (5.5) provides a formula for the coarse-grained holding rates, denoted by $\boldsymbol{\lambda}_N(j)$:

$$
\boldsymbol{\lambda}_N(k) := \sum_{j \neq k} \boldsymbol{r}_N(k,j) = \frac{\theta_N}{\pi_N(\mathcal{E}^k)} \sum_{\zeta \in \mathcal{E}^k} \pi_N(\zeta)\,\lambda(\zeta)\,\mathbb{P}_\zeta\big[\,H_{\breve{\mathcal{E}}^k} < H_{\mathcal{E}^k}^+\,\big]. \quad (6.1)
$$

The expression on the right-hand side corresponds to the capacity between $\mathcal{E}^k$ and $\breve{\mathcal{E}}^k$. Therefore,

$$
\pi_N(\mathcal{E}^j)\,\boldsymbol{\lambda}_N(j) = \theta_N \, \mathrm{cap}_N(\mathcal{E}^j, \breve{\mathcal{E}}^j)\,, \quad j \in S\,. \quad (6.2)
$$

Remark 6.1. *Equation* (6.2) *provides a formula for the magnitude of the scaling parameter θ_N. To derive a non-trivial limit for the coarse-grained model $\boldsymbol{X}_N^T$, time has to be rescaled by the inverse of the capacity between the sets $\mathcal{E}^j$ and $\breve{\mathcal{E}}^j$:*

$$\theta_N \;\approx\; \frac{\pi_N(\mathcal{E}^j)}{\mathrm{cap}_N(\mathcal{E}^j, \breve{\mathcal{E}}^j)} \;.$$

The asymptotic behavior of the coarse-grained holding rates can be computed through formula (6.2) provided one can estimate the capacities and the measures of the valleys. Once this has been done, to compute the jump rates, it remains to estimate the jump probabilities.

Recall from Section A the definition of a collapsed chain. Fix $j \in S$, and denote by $\eta^{C,j}(t)$ the Markov chain obtained from the chain $\eta_N(t)$ by collapsing the valley $\mathcal{E}^j$ to a point, denoted by $\mathfrak{j}$. The chain $\eta^{C,j}(t)$ takes value in $E_N^{C,j} := (E_N \setminus \mathcal{E}^j) \cup \{\mathfrak{j}\}$.

Let $\mathbb{P}_\eta^{C,j}$, $\eta \in E_N^{C,j}$, be the probability measure on $D([0,\infty), E_N^{C,j})$ induced by the collapsed process $\eta^{C,j}(t)$ starting from η. Expectation with respect to $\mathbb{P}_\eta^{C,j}$ is represented by $\mathbb{E}_\eta^{C,j}$. By the last formula of the proof of [19, Proposition 3.4], for any $k \in S$, $k \neq j$,

$$\boldsymbol{p}_N(j,k) \;:=\; \frac{\boldsymbol{r}_N(j,k)}{\boldsymbol{\lambda}_N(j)} \;=\; \mathbb{P}_{\mathfrak{j}}^{C,j}\big[\, H_{\mathcal{E}^k} < H_{\breve{\mathcal{E}}^{j,k}}\,\big]\,, \quad \text{where} \quad \breve{\mathcal{E}}^{j,k} \;:=\; \bigcup_{\ell \in S \setminus \{j,k\}} \mathcal{E}_N^\ell \;.$$

Denote by $\boldsymbol{P}_j$ the probability measure on $D([0,\infty), S)$ induced by the reduced model $\boldsymbol{X}(t)$ starting from j. We present below a set of sufficient conditions which ensure that $\boldsymbol{p}_N(j,k)$ converges to $\boldsymbol{P}_j[H_k < H_{S \setminus \{j,k\}}]$. This approach has been developed and gradually refined in [85, 92, 117], and it is based on the premise that the capacities can be calculated through the Thomson and the Dirichlet principles.

Denote by $L^2(\pi_N)$ the space of square-summable functions $f : E_N \to \mathbb{R}$ endowed with the scalar product $\langle \cdot, \cdot \rangle_{\pi_N}$ given by

$$\langle\, f\,, g\,\rangle_{\pi_N} \;:=\; \sum_{\eta \in E_N} f(\eta)\, g(\eta)\, \pi_N(\eta) \;.$$

We assume that the generator $\mathcal{L}_N$ of the Markov chain $\eta_N(t)$ satisfies a *sector condition* with a constant C_0 independent of N: For every $f, g \in L^2(\pi_N)$,

$$\langle\, \mathcal{L}_N f\,, g\,\rangle_{\pi_N}^2 \;\leq\; C_0\, \langle\, (-\mathcal{L}_N)f\,, f\,\rangle_{\pi_N} \langle\, (-\mathcal{L}_N)g\,, g\,\rangle_{\pi_N} \;. \qquad (6.3)$$

Suppose that for fixed $j, k \in S$, $k \neq j$,

$$\lim_{N \to \infty} \theta_N\, \mathrm{cap}_N(\mathcal{E}^k, \breve{\mathcal{E}}^{j,k}) \;=\; \mathrm{cap}_S(k, S \setminus \{j,k\})\,,$$
$$\lim_{N \to \infty} \theta_N\, \mathrm{cap}_N(\mathcal{E}^j, \breve{\mathcal{E}}^j) \;=\; \mathrm{cap}_S(j, S \setminus \{j\})\,, \qquad (6.4)$$

where $\mathrm{cap}_S(A,B)$, $A, B \subset S$, represents the capacity with respect to the reduced model $\boldsymbol{X}(t)$.

We also assume that the capacities for the collapsed process $\eta^{C,j}(t)$ can be calculated: Denote by $\mathrm{cap}_N^{C,j}(\mathcal{A}, \mathcal{B})$ the capacity between $\mathcal{A}, \mathcal{B} \subset E_N^{C,j}$, $\mathcal{A} \cap \mathcal{B} = \varnothing$ induced by the collapsed process $\eta^{C,j}(t)$. We assume that the limit of the capacity $\mathrm{cap}_N^{C,j}(\mathcal{E}^k, \breve{\mathcal{E}}^{j,k})$ coincides with $\mathrm{cap}_S(k, S \setminus \{j,k\})$:

$$\lim_{N \to \infty} \theta_N \, \mathrm{cap}_N^{C,j}(\mathcal{E}^k, \breve{\mathcal{E}}^{j,k}) \;=\; \mathrm{cap}_S(k, S \setminus \{j,k\}) \,. \tag{6.5}$$

The computation of the capacities requires test flows or test functions which approximate the optimal ones in the variational principles. It is thus implicitly assumed in hypotheses (6.4) and (6.5) that explicit expressions for such flows or functions are available. We assume below that there exists a sequence of functions $V_{j,k}^N : \mathcal{E}_N \to [0,1]$ close to the *equilibrium potential* $h_{\mathcal{E}^k, \breve{\mathcal{E}}^{j,k}}^N$, given by

$$h_{\mathcal{E}^k, \breve{\mathcal{E}}^{j,k}}^N(\eta) \;:=\; \mathbb{P}_\eta^N \big[\, H_{\mathcal{E}^k} < H_{\breve{\mathcal{E}}^{j,k}} \,\big] \,,$$

in the sense that

$$\lim_{N \to \infty} \theta_N \, D_N(V_{j,k}^N) \;=\; \lim_{N \to \infty} \theta_N \, D_N\big(h_{\mathcal{E}^k, \breve{\mathcal{E}}^{j,k}}^N\big) \;=\; \mathrm{cap}_S(k, S \setminus \{j,k\}) \,, \tag{6.6}$$

where $D_N(f)$ stands for the Dirichlet form of f:

$$D_N(f) \;:=\; \langle\, (-\mathcal{L}_N)\, f \,,\, f \,\rangle_{\pi_N} \,.$$

The last identity in (6.6) follows from the fact, proved in (B.7), that $D_N(h_{\mathcal{E}^k, \breve{\mathcal{E}}^{j,k}}^N) = \mathrm{cap}_N(\mathcal{E}^k, \breve{\mathcal{E}}^{j,k})$ and from assumption (6.4).

We assume, furthermore, that $V_{j,k}^N$ is constant in each valley $\mathcal{E}_N^\ell$:

$$V_{j,k}^N(\eta) \;=\; \sum_{\ell \in S} \mathbf{P}_\ell[H_k < H_{S \setminus \{j,k\}}] \, \chi_{\mathcal{E}^\ell}(\eta) \quad \text{for all} \quad \eta \in \mathcal{E}_N \,. \tag{6.7}$$

Hence, $V_{j,k}^N$ is equal to 1, 0 in $\mathcal{E}_N^k$, $\breve{\mathcal{E}}_N^{j,k}$, respectively, while on $\mathcal{E}_N^j$ it is given by the probability appearing in (6.7).

Finally, as $V_{j,k}^N$ approximates $h_{\mathcal{E}^k, \breve{\mathcal{E}}^{j,k}}^N$, which is harmonic on $\Delta_N \cup \mathcal{E}^j$, it is also reasonable to require $\mathcal{L}_N V_{j,k}^N$ to be small in these sets. We assume that

$$\lim_{N \to \infty} \theta_N \sum_{\eta \in \Delta_N} \big|(\mathcal{L}_N V_{j,k}^N)(\eta)\big| \pi_N(\eta) = \lim_{N \to \infty} \theta_N \Big| \sum_{\eta \in \mathcal{E}^j} (\mathcal{L}_N V_{j,k}^N)(\eta) \pi_N(\eta) \Big| = 0. \tag{6.8}$$

Proposition 6.2. *Fix j, $k \in S$, $k \neq j$, and assume that conditions (6.3)–(6.8) are in force. Then,*

$$\lim_{N \to \infty} \boldsymbol{p}_N(j,k) \;=\; \mathbf{P}_j[H_k < H_{S \setminus \{j,k\}}] \,.$$

The proof of this proposition is divided in several lemmata. Since $V_{j,k}^N$ is constant on $\mathcal{E}^j$, we may collapse it to a function defined on $E_N^{C,j}$. Recall from

(6.7) the value of $V_{j,k}^N$ at $\mathcal{E}^j$ and let $V_{j,k}^{C,j} : E_N^{C,j} \to [0,1]$ be given by

$$V_{j,k}^{C,j}(\mathfrak{j}) \;=\; \mathbf{P}_j[H_k < H_{S \setminus \{j,k\}}]\,, \quad V_{j,k}^{C,j}(\eta) \;=\; V_{j,k}^N(\eta)\,, \quad \eta \in E_N^{C,j} \setminus \{\mathfrak{j}\}\,.$$

The dependence of $V_{j,k}^{C,j}$ on N has been omitted.

Let $\mathcal{L}_N^{C,j}$ be the generator of the collapsed process $\eta^{C,j}(t)$. For $\mathcal{A},\, \mathcal{B} \subset E_N^{C,j}$, $\mathcal{A} \cap \mathcal{B} = \varnothing$, denote by $h_{\mathcal{A},\mathcal{B}}^{(\mathfrak{j})}$ the solution of the boundary value elliptic problem

$$\left\{ \begin{array}{l} (\, \mathcal{L}_N^{C,j}\, h\,)\,(\eta) \;=\; 0\,, \;\; \eta \notin \mathcal{A} \cup \mathcal{B}\,, \\ h(\eta) \;=\; \chi_{\mathcal{A}}(\eta)\,, \;\; \eta \in \mathcal{A} \cup \mathcal{B}\,. \end{array} \right.$$

Denote by $h_{j,k} : E_N \to \mathbb{R}$ the lifting of the function $h_{\mathcal{E}^k,\breve{\mathcal{E}}^{j,k}}^{(\mathfrak{j})}$:

$$h_{j,k}(\eta) \;=\; h_{\mathcal{E}^k,\breve{\mathcal{E}}^{j,k}}^{(\mathfrak{j})}(\mathfrak{j})\,, \quad \eta \in \mathcal{E}^j\,, \qquad h_{j,k}(\eta) \;=\; h_{\mathcal{E}^k,\breve{\mathcal{E}}^{j,k}}^{(\mathfrak{j})}(\eta)\,, \quad \eta \in E_N \setminus \mathcal{E}^j\,.$$

Note that the function $h_{j,k}$ is constant and equal to $\boldsymbol{p}_N(j,k)$ on the set $\mathcal{E}^j$.

Lemma 6.3. *We claim that*

$$\lim_{N \to \infty} \theta_N\, D_N\big(\, h_{j,k} - V_{j,k}^N\,\big) \;=\; 0\,.$$

Proof. Rewrite $D_N(h-V)$ as $\langle\, h-V\,,\, (-\mathcal{L}_N\,[h-V]\,)\,\rangle_{\pi_N}$ and compute separately the limit of the four terms.

By equation (A.15), $D_N(h_{j,k}) = D_N^{C,j}(h_{\mathcal{E}^k,\breve{\mathcal{E}}^{j,k}}^{(\mathfrak{j})})$, where $D_N^{C,j}$ represents the Dirichlet form associated to the collapsed process. By (B.7), $D_N^{C,j}(h_{\mathcal{E}^k,\breve{\mathcal{E}}^{j,k}}^{(\mathfrak{j})}) = \mathrm{cap}_N^{C,j}(\mathcal{E}^k, \breve{\mathcal{E}}^{j,k})$. Hence, by assumption (6.4),

$$\lim_{N \to \infty} \theta_N\, D_N(h_{j,k}) \;=\; \mathrm{cap}_S(k, S \setminus \{j,k\})\,.$$

By assumption (6.6), the same result holds for $V_{j,k}$ in place of $h_{j,k}$.

It remains to examine the cross terms. By (A.15),

$$\langle\, V_{j,k}\,,\, (-\mathcal{L}_N\, h_{j,k})\,\rangle_{\pi_N} \;=\; \langle\, V_{j,k}^{C,j}\,,\, (-\mathcal{L}_N^{C,j}\, h_{\mathcal{E}^k,\breve{\mathcal{E}}^{j,k}}^{(\mathfrak{j})})\,\rangle_{\pi_N^{C,j}}\,,$$

where $\pi_N^{C,j}$ stands for the stationary measure π_N collapsed at $\mathcal{E}^j$. Since $h_{\mathcal{E}^k,\breve{\mathcal{E}}^{j,k}}^{(\mathfrak{j})}$ is harmonic on $\Delta_N \cup \{\mathfrak{j}\}$, and since $V_{j,k}^{C,j}$ vanishes on $\breve{\mathcal{E}}^{j,k}$ and coincides with $h_{\mathcal{E}^k,\breve{\mathcal{E}}^{j,k}}^{(\mathfrak{j})}$ on $\mathcal{E}^k$, the last expression is equal to

$$\sum_{\eta \in \mathcal{E}^k} V_{j,k}^{C,j}(\eta)\,(-\mathcal{L}_N^{C,j}\, h_{\mathcal{E}^k,\breve{\mathcal{E}}^{j,k}}^{(\mathfrak{j})})\,(\eta)\, \pi_N^{C,j}(\eta)$$

$$= \sum_{\eta \in \mathcal{E}^k} h_{\mathcal{E}^k,\breve{\mathcal{E}}^{j,k}}^{(\mathfrak{j})}(\eta)\,(-\mathcal{L}_N^{C,j}\, h_{\mathcal{E}^k,\breve{\mathcal{E}}^{j,k}}^{(\mathfrak{j})})\,(\eta)\, \pi_N^{C,j}(\eta)\,.$$

Using again the harmonicity of $h^{(j)}_{\mathcal{E}^k, \check{\mathcal{E}}^{j,k}}$ on $\Delta_N \cup \{j\}$, and the fact that it vanishes on $\check{\mathcal{E}}^{j,k}$, we may extend the sum to the entire set $E_N^{C,j}$ and conclude, as at the beginning of the proof, that

$$\lim_{N \to \infty} \theta_N \langle V_{j,k}, (-\mathcal{L}_N h_{j,k}) \rangle_{\pi_N} = \mathrm{cap}_S(k, S \setminus \{j, k\}) .$$

Similarly, since $h_{j,k}$ is equal to $V_{j,k}$ on $\mathcal{E}^k \cup \check{\mathcal{E}}^{j,k}$,

$$\langle h_{j,k}, (-\mathcal{L}_N V_{j,k}) \rangle_{\pi_N} = \sum_{\eta \in \mathcal{E}^k} V_{j,k}(\eta) (-\mathcal{L}_N V_{j,k})(\eta) \pi_N(\eta)$$
$$+ \sum_{\eta \in \Delta_N \cup \mathcal{E}^j} h_{j,k}(\eta) (-\mathcal{L}_N V_{j,k})(\eta) \pi_N(\eta) .$$

Since $V_{j,k}$ vanishes on $\check{\mathcal{E}}^{j,k}$, the first term on the righ-hand side is equal to

$$D_N(V_{j,k}) - \sum_{\eta \in \Delta_N \cup \mathcal{E}^j} V_{j,k}(\eta) (-\mathcal{L}_N V_{j,k})(\eta) \pi_N(\eta) .$$

Therefore,

$$\langle h_{j,k}, (-\mathcal{L}_N V_{j,k}) \rangle_{\pi_N} = D_N(V_{j,k})$$
$$+ \sum_{\eta \in \Delta_N \cup \mathcal{E}^j} \{ h_{j,k}(\eta) - V_{j,k}(\eta) \} (-\mathcal{L}_N V_{j,k})(\eta) \pi_N(\eta) .$$

Since $h_{j,k}$ and $V_{j,k}$ are constant in $\mathcal{E}^j$, non-negative and bounded by 1, the absolute value of the second term on the right-hand side is less than or equal to

$$\sum_{\eta \in \Delta_N} \left| (\mathcal{L}_N V_{j,k})(\eta) \right| \pi_N(\eta) + \left| \sum_{\eta \in \mathcal{E}^j} (\mathcal{L}_N V_{j,k})(\eta) \pi_N(\eta) \right| .$$

By condition (6.8), this expression multiplied by θ_N converges to 0 as $N \to \infty$. Thus, by (6.6),

$$\lim_{N \to \infty} \theta_N \langle h_{j,k}, (-\mathcal{L}_N V_{j,k}) \rangle_{\pi_N} = \mathrm{cap}_S(k, S \setminus \{j, k\}) .$$

Putting together all previous estimates yields the assertion. $\qquad\square$

Fix two non-empty subsets $\mathcal{A}, \mathcal{B}$ of E_N such that $\mathcal{A} \cap \mathcal{B} = \varnothing$, Recall from Section B that we represent by $\mathfrak{C}_{1,0}(\mathcal{A}, \mathcal{B})$ the space of functions $f : E_N \to [0, 1]$ which are equal to 1 on $\mathcal{A}$ and 0 on $\mathcal{B}$. Let $f_{j,k}^N = h_{j,k} - V_{j,k}$, and note that this function is constant on each valley $\mathcal{E}^\ell$.

Lemma 6.4. *Let $f_{j,k}(\mathcal{E}^j)$ be the value of $f_{j,k}$ at $\mathcal{E}^j$. Then, $\lim_{N \to \infty} f_{j,k}(\mathcal{E}^j) = 0$.*

Proof. The function $f_{j,k}$ vanishes on $\mathcal{E}^k \cup \check{\mathcal{E}}^{j,k} = \check{\mathcal{E}}^j$, and it is constant on $\mathcal{E}^j$. Hence, if $f_{j,k}(\mathcal{E}^j) \neq 0$, the function F defined as $F(\eta) = f_{j,k}(\eta)/f_{j,k}(\mathcal{E}^j)$ belongs to $\mathfrak{C}_{1,0}(\mathcal{E}^j, \check{\mathcal{E}}^j)$.

Suppose that $f_{j,k}(\mathcal{E}^j) \neq 0$. On the one hand, by Lemma 6.3, $\theta_N f_{j,k}(\mathcal{E}^j)^2 D_N(F) = \theta_N D_N(f_{j,k}) \to 0$. On the other hand, since F belongs to $\mathfrak{C}_{1,0}(\mathcal{E}^j, \breve{\mathcal{E}}^j)$, by (B.14), $D_N(F) \geq \mathrm{cap}_N^s(\mathcal{E}^j, \breve{\mathcal{E}}^j)$, where $\mathrm{cap}_N^s(\mathcal{E}^j, \breve{\mathcal{E}}^j)$ represents the capacity associated to the symmetric dynamics. By the sector condition, stated in assumption (6.3), and Lemma B.15, this symmetric capacity is bounded below by $c_0 \mathrm{cap}_N(\mathcal{E}^j, \breve{\mathcal{E}}^j)$, where $c_0 = 1/C_0 > 0$. Hence, by (6.4), $\liminf_{N\to\infty} \theta_N D_N(F_N) \geq c_0 \mathrm{cap}_S(j, S \setminus \{j\}) > 0$, which proves the assertion of the lemma. $\qquad \square$

Proof of Proposition 6.2. By definition (6.7) of $V_{j,k}$, $V_{j,k}(\mathcal{E}^j) = \mathbf{P}_j[H_k < H_{S\setminus\{j,k\}}]$. The claim of the proposition follows from Lemma 6.4 and the fact that

$$h_{j,k}(\mathcal{E}^j) = h^{(\mathrm{j})}_{\mathcal{E}^k, \breve{\mathcal{E}}_{j,k}}(\mathrm{j}) = \boldsymbol{p}_N(j,k) . \qquad \square$$

Remark 6.5. *Assumption* (6.3) *can be replaced by the hypothesis that*

$$\liminf_{N\to\infty} \theta_N \mathrm{cap}_N^s(\mathcal{E}^j, \breve{\mathcal{E}}^j) > 0 . \tag{6.9}$$

Proof. We only used the sector condition, assumption (6.3), in the proof of Lemma 6.4 to guarantee that $\theta_N D_N(F)$ is bounded below by a strictly positive constant. Since $D_N(F) \geq \mathrm{cap}_N^s(\mathcal{E}^j, \breve{\mathcal{E}}^j)$, by (6.9), $\liminf_{N\to\infty} \theta_N D_N(F) > c_0$, as needed. $\qquad \square$

7. The negligible set $\boldsymbol{\Delta}_N$

We provide in this section sufficient conditions for assumption (LP2) or (T2) to hold. Recall from (5.5) that $\boldsymbol{r}_N(k,j)$ represents the coarse-grained jump rates. Assume that they converge: For all $j \neq k$, there exists $\boldsymbol{r}(k,j) \in [0,\infty)$ such that

$$\lim_{N\to\infty} \boldsymbol{r}_N(k,j) = \boldsymbol{r}(k,j) . \tag{7.1}$$

Recall that we represent by $\boldsymbol{X}(t)$ the reduced model, the S-valued Markov chain whose jump rates are given by $\boldsymbol{r}(k,j)$. Denote by $A \subset S$ the subset of S formed by the points which are absorbing for the reduced model $\boldsymbol{X}(t)$. Next result is Theorem 2.7 in [15] and Theorem 2.1 in [19].

Theorem 7.1. *Assume that conditions* (5.4) *and* (7.1) *are in force. Assume, furthermore, that for all $k \in A$, $t > 0$,*

$$\lim_{N\to\infty} \max_{\eta\in\mathcal{E}_N^k} \mathbb{E}_\eta^N \left[\int_0^t \chi_{\Delta_N}\big(\eta_N(s\theta_N)\big) \, ds \right] = 0 ; \tag{7.2}$$

and that for all $j \notin A$

$$\lim_{N\to\infty} \frac{\pi_N(\Delta_N)}{\pi_N(\mathcal{E}_N^j)} = 0 . \tag{7.3}$$

Then, property (LP2) is in force.

Remark 7.2. *In the previous theorem, we may replace conditions* (7.2), (7.3) *by the assumption*

$$\lim_{N\to\infty} \frac{1}{\sum_{k\neq j} \boldsymbol{r}_N(j,k)} \frac{\pi_N(\Delta_N)}{\pi_N(\mathcal{E}_N^j)} \;=\; 0$$

for all $j \in S$.

In some spin dynamics, the valleys are formed by few configurations and the following simple argument applies.

Lemma 7.3. *Assume that*

$$\lim_{N\to\infty} \max_{\eta\in\mathcal{E}_N} \frac{\pi_N(\Delta_N)}{\pi_N(\eta)} \;=\; 0 \;.$$

Then, condition (LP2) is in force.

Proof. Fix $t > 0$. Clearly, dividing and multiplying by $\pi_N(\eta)$,

$$\mathbb{E}_\eta^N\Big[\int_0^t \chi_{\Delta_N}\big(\eta_N(s\theta_N)\big)\, ds \Big] \;\leq\; \frac{1}{\pi_N(\eta)}\, \mathbb{E}_{\pi_N}^N\Big[\int_0^t \chi_{\Delta_N}\big(\eta_N(s\theta_N)\big)\, ds \Big] \;.$$

Since π_N is the stationary state, the previous expression is equal to $t\,\pi_N(\Delta_N)/\pi_N(\eta)$, which proves the lemma. $\qquad\square$

8. The Poisson equation

We present here an alternative method to prove uniqueness of limit points of the sequence of measures $\mathbf{P}^N$ introduced in Section 4. It relies on asymptotic properties of the solutions of Poisson equations.

Assume that we are able to foretell the dynamics of the reduced model, and denote by L its generator. Fix a function $F\colon S \to \mathbb{R}$, and let $G = LF$. Denote by f, $g\colon E_N \to \mathbb{R}$ the function given by

$$f \;=\; \sum_{k\in S} F(k)\,\chi_{\mathcal{E}_N^k}\,, \quad g \;=\; \sum_{k\in S} G(k)\,\chi_{\mathcal{E}_N^k}\,. \qquad (8.1)$$

The functions f, g are constant on each valley $\mathcal{E}_N^\ell$ and vanish at Δ_N. The method presented below relies on the assumption that the solution f_N of the Poisson equation

$$\theta_N\,\mathcal{L}_N\,f_N \;=\; g$$

is almost constant on each set $\mathcal{E}_N^k$. A solution of this equation exists only if g has zero-mean with respect to π_N, which is not necessarily the case. Therefore, we need first to turn g into a zero-mean function and then to consider the solution of the Poisson equation. This is the content of conditions (A1), (A2).

Assume that there exists a sequence of function $g_N \colon E_N \to \mathbb{R}$ such that

(A1) g_N has zero-mean with respect to π_N, vanishes on Δ_N and converges to g uniformly on $\mathcal{E}_N$;

(A2) Denote by f_N the unique solution of the Poisson equation

$$\theta_N \, \mathcal{L}_N f \ = \ g_N \qquad (8.2)$$

in E_N. There exists a finite constant C_0 such that

$$\sup_{N \geq 1} \max_{\eta \in E_N} |f_N(\eta)| \ \leq \ C_0 \,, \quad \text{and} \quad \lim_{N \to \infty} \max_{\eta \in \mathcal{E}_N} \big| \, f_N(\eta) - f(\eta) \, \big| \ = \ 0 \,.$$

The natural candidate for g_N in conditions (A1) and (A2) is the function g itself, but it does not have zero-mean. To fulfill this condition, denote by π the stationary state of the reduced model. We expect $\pi_N(\mathcal{E}_N^k)$ to converge to $\pi(k)$. Hence,

$$\lim_{N \to \infty} E_{\pi_N}[g] \ = \ \lim_{N \to \infty} \sum_{k \in S} G(k) \, \pi_N(\mathcal{E}_N^k) \ = \ \sum_{k \in S} (LF)(k) \, \pi(k) \ = \ 0 \,.$$

A reasonable candidate for g_N is thus $g - \epsilon_N \, \chi_{\mathcal{E}_N^1}$, where $\epsilon_N = E_{\pi_N}[g]/\pi_N(\mathcal{E}_N^1)$ vanishes as $N \to \infty$ [if $\pi_N(\mathcal{E}_N^1) \to \pi(1) > 0$].

Properties (A1), (A2) have been proved in [57, 118] for elliptic operators on $\mathbb{R}^d$ of the form $\mathcal{L}_N f \ = \ e^{NV} \nabla \cdot (e^{-NV} a \nabla f)$, in [94] for one-dimensional diffusions with periodic boundary conditions, and in [106] for condensing zero-range processes.

The main result of this section, Theorem 8.2 below, asserts that conditions (A1), (A2) guarantee uniqueness of limit points of the sequence $\mathbf{P}^N$. The proof of this result requires some preparation.

Let $\mathbb{Q}_\eta^N$, $\eta \in E_N$, be the probability measure on $D([0,\infty), E_N)$ induced by the speeded-up process $\xi^N(t) := \eta_N(t\theta_N)$ starting from η. Keep in mind that the generator of this process is $\theta_N \, \mathcal{L}_N$. Denote by $(\mathcal{F}_t^o : t \geq 0)$ the σ-algebra of subsets of $D([0,\infty), E_N)$ generated by $\{\eta(s) : 0 \leq s \leq t\}$, where $\eta(s)$ represents the coordinate process. Fix $\eta \in E_N$ and denote by $\{\mathcal{F}_t^\eta : t \geq 0\}$ the usual augmentation of $\{\mathcal{F}_t^o : t \geq 0\}$ with respect to $\mathbb{Q}_\eta^N$. We refer to Section III.9 of [115] for a precise definition. The advantage of $\mathcal{F}_t^\eta$ with respect to $\mathcal{F}_t^o$ is that it is right-continuous: $\mathcal{F}_t^\eta = \cap_{s>t} \mathcal{F}_s^\eta$.

Recall from (2.4) the definition of the time change $S_{\mathcal{E}_N}(t)$ associated to the additive functional $T_{\mathcal{E}_N}(t)$. Clearly, for all $r \geq 0$, $t \geq 0$,

$$\{S_{\mathcal{E}_N}(r) \geq t\} \ = \ \{T_{\mathcal{E}_N}(t) \leq r\} \,. \qquad (8.3)$$

Lemma 8.1. *For each $t \geq 0$ and $\eta \in E_N$, $S_{\mathcal{E}_N}(t)$ is a stopping time with respect to the filtration $(\mathcal{F}_t^\eta : t \geq 0)$.*

Proof. Fix $t \geq 0$, $r \geq 0$ and $\eta \in E_N$. By (8.3),

$$\{S_{\mathcal{E}_N}(t) \leq r\} \ = \ \bigcap_q \{S_{\mathcal{E}_N}(t) < r + q\} \ = \ \bigcap_q \{T_{\mathcal{E}_N}(r+q) > t\} \,,$$

where the intersection is carried out over all $q \in (0, \infty) \cap \mathbb{Q}$. By definition of $T_{\mathcal{E}_N}$, $\{T_{\mathcal{E}_N}(r + q) > t\}$ belongs to $\mathcal{F}_{r+q}^\eta$. Hence, as the filtration is right-continuous, $\{S_{\mathcal{E}_N}(t) \leq r\} \in \cap_q \mathcal{F}_{r+q}^\eta = \mathcal{F}_r^\eta$, which proves the lemma.   □

Let $(\mathcal{G}_t^{N,\eta} : t \geq 0)$ be the filtration given by $\mathcal{G}_t^{N,\eta} = \mathcal{F}_{S_{\mathcal{E}_N}(t)}^\eta$, and denote by $\xi^{\mathcal{E}_N}(t)$ the trace of the coordinate process $\eta(t)$ on $\mathcal{E}_N$: $\xi^{\mathcal{E}_N}(t) = \eta(S_{\mathcal{E}_N}(t))$. Clearly, the process $\xi^{\mathcal{E}_N}(t)$ is adapted to the filtration $(\mathcal{G}_t^{N,\eta})$. Moreover, as the coordinate process corresponds to the distribution of $\xi^N(t)$, $\xi^{\mathcal{E}_N}(t)$ corresponds to the trace of the speeded-up process $\xi^N(t)$ on $\mathcal{E}_N$.

It is easy to check that we may commute the trace operation with the acceleration of the process:

$$\eta^{\mathcal{E}_N}(t\theta_N) = \xi^{\mathcal{E}_N}(t) .$$

On the left-hand side, we first computed the trace of the chain $\eta_N(t)$ on $\mathcal{E}_N$ and then accelerated it by θ_N, while on the right-hand side we first speeded-up the chain $\eta_N(t)$ by θ_N and then computed the trace of the result on $\mathcal{E}_N$. In particular, the process $\boldsymbol{X}_N^T(t)$, introduced in assumption (T1), corresponds to the projection of $\xi^{\mathcal{E}_N}(t)$ on S through Ψ_N:

$$\boldsymbol{X}_N^T(t) = \Psi_N(\xi^{\mathcal{E}_N}(t)) . \tag{8.4}$$

Moreover, the measure $\mathbf{P}^N$ on $D([0, \infty), S)$ represents the distribution of the process $\Psi_N(\xi^{\mathcal{E}_N}(t))$. We may now state the main result of this section.

Theorem 8.2. *Fix $k \in S$ and a sequence $\eta_N \in \mathcal{E}_N^k$. Assume that conditions (A1) and (A2) are in force for every function $F : S \to \mathbb{R}$. Then, every limit point $\mathbf{P}$ of the sequence $\mathbf{P}^N$ such that*

$$\mathbf{P}\big[\, X(t-) = X(t) \,\big] \quad \text{for all } t > 0 . \tag{8.5}$$

solves the (L, δ_k) martingale problem.

Proof. Fix a function $F : S \to \mathbb{R}$. Let $f_N : E_N \to \mathbb{R}$ be the function given by assumption (A2). Then,

$$M_N(t) = f_N(\eta(t)) - f_N(\eta(0)) - \int_0^t \theta_N \, (\mathcal{L}_N f_N)(\eta(s)) \, ds$$

$$= f_N(\eta(t)) - f_N(\eta(0)) - \int_0^t g_N(\eta(s)) \, ds$$

is a martingale in $(D([0, \infty), E_N), (\mathcal{F}_t^{\eta_N}), \mathbb{Q}_{\eta_N}^N)$.

Since $\{S_{\mathcal{E}_N}(t) : t \geq 0\}$ are stopping times with respect to the filtration $(\mathcal{F}_t^{\eta_N})$, $M_N(S_{\mathcal{E}_N}(t))$ is a martingale with respect to $(\mathcal{F}_{S_{\mathcal{E}_N}(t)}^{\eta_N}) = (\mathcal{G}_t^{N,\eta_N})$. Hence, by definition of the trace process $\xi^{\mathcal{E}_N}(t)$,

$$\widehat{M}_N(t) := M_N(S_{\mathcal{E}_N}(t)) = f_N(\xi^{\mathcal{E}_N}(t)) - f_N(\xi^{\mathcal{E}_N}(0)) - \int_0^{S_{\mathcal{E}_N}(t)} g_N(\eta(s)) \, ds$$

is a martingale with respect to the filtration $\mathcal{G}_t^{N,\eta_N}$. Since g_N vanishes on Δ_N, we may insert in the integral the indicator function of the set $\mathcal{E}_N$. Then, a change of variables yields that this integral is equal to

$$\int_0^{S_{\mathcal{E}_N}(t)} g_N(\eta(s)) \, \chi_{\mathcal{E}_N}(\eta(s)) \, ds \;=\; \int_0^t g_N(\eta(S_{\mathcal{E}_N}(s))) \, ds \;.$$

Therefore,

$$\widehat{M}_N(t) \;=\; f_N(\xi^{\mathcal{E}_N}(t)) \;-\; f_N(\xi^{\mathcal{E}_N}(0)) \;-\; \int_0^t g_N(\xi^{\mathcal{E}_N}(s)) \, ds$$

is a $\{\mathcal{G}_t^{N,\eta_N}\}$-martingale.

By (A1) and (A2), g_N, resp. f_N, converge to g, resp. f, uniformly in $\mathcal{E}^N$ as $N \to 0$. Hence, since $\xi^{\mathcal{E}_N}(s) \in \mathcal{E}^N$ for all $s \geq 0$, we may replace in the previous equation g_N, f_N by g, f, respectively, at a cost which vanishes as $N \to \infty$. Therefore,

$$\widehat{M}_N(t) \;=\; f(\xi^{\mathcal{E}_N}(t)) \;-\; f(\xi^{\mathcal{E}_N}(0)) \;-\; \int_0^t g(\xi^{\mathcal{E}_N}(s)) \, ds \;+\; o_N(1)$$

is a $\{\mathcal{G}_t^{N,\eta_N}\}$-martingale.

Since f and g are constant on each set $\mathcal{E}_N^k$, by (8.1) and (8.4), $f(\xi^{\mathcal{E}_N}(t)) = F(\Psi_N(\xi^{\mathcal{E}_N}(t))) = F(\boldsymbol{X}_N^T(t))$, $g(\xi^{\mathcal{E}_N}(s)) = G(\Psi_N(\xi^{\mathcal{E}_N}(s))) = G(\boldsymbol{X}_N^T(s))$, and

$$F(\boldsymbol{X}_N^T(t)) \;-\; F(\boldsymbol{X}_N^T(0)) \;-\; \int_0^t (LF)(\boldsymbol{X}_N^T(s)) \, ds \;+\; o_N(1)$$

is a martingale because $G = LF$.

Since $\mathbf{P}^N$ corresponds to the distribution of $\boldsymbol{X}_N^T$,

$$\widehat{M}(t) \;=\; F(X(t)) \;-\; F(X(0)) \;-\; \int_0^t (LF)(X(s)) \, ds$$

is a martingale under $\mathbf{P}^N$ up to a small error. Let $\mathbf{P}$ be a limit point of the sequence $\mathbf{P}^N$ satisfying (8.5), and assume, without loss of generality, that $\mathbf{P}^N$ converges to $\mathbf{P}$. By (8.5), the one-dimensional projections are continuous, and we may pass to the limit to obtain that $\widehat{M}(t)$ is a martingale under $\mathbf{P}$.

On the other hand, as $\eta_N \in \mathcal{E}_N^k$, $\mathbf{P}^N[X(0) = k] = 1$ for all N, so that $\mathbf{P}[X(0) = k] = 1$. This proves that any limit point of the sequence $\mathbf{P}^N$ satisfying (8.5) is a solution of the (L, δ_k) martingale problem. $\square$

9. Local ergodic theorem in L^2

It is not clear whether the scheme presented in the previous section can be applied to a large class of dynamics. The proof of condition (A2) is unclear even for the simple example of Section 2.

The method presented in Sections 4–6 has also a drawback. As the function $f = \sum_{k \in S} F(k) \chi_{\mathcal{E}_N^k}$ has a sharp interface, the jump rates $R_N^{(k)}$ which appear in the computation of $\mathcal{L}_{\mathcal{E}_N} f$, are singular functions, vanishing at the interior of the valleys and taking large values at the boundary. This lack of smoothness turns the proof of the local ergodic theorem more demanding.

Following [20], we propose below an alternative approach, in which we replace the indicator function $\chi_{\mathcal{E}_N^k}$ by "smooth" approximations obtained by solving the resolvent equation

$$\left(I - \gamma_N \mathcal{L}_{\mathcal{E}_N} \right) f = \chi_{\mathcal{E}_N^k} , \tag{9.1}$$

where $\mathcal{L}_{\mathcal{E}_N}$ represents the generator of the trace process $\eta^{\mathcal{E}_N}(t)$, I the identity and γ_N a suitable sequence of positive numbers.

The resolvent equation (9.1) has a unique solution, denoted by u_N^k. Equation (9.12) provides a stochastic representation of the solution, different from the usual one given in terms of a time integral. This guarantees existence. Uniqueness can be proven as follows. Let u_1, u_2 be two solutions, and set $w = u_1 - u_2$. The function w solves (9.1) with a right-hand side equal to 0. Multiply both sides of the equation by w and integrate with respect to $\pi_{\mathcal{E}}$ to get that $w = 0$ because $\langle \mathcal{L}_{\mathcal{E}_N} f , f \rangle_{\pi_{\mathcal{E}}} \leq 0$ for all functions $f : \mathcal{E}_N \to \mathbb{R}$.

Note that $\gamma_N \mathcal{L}_{\mathcal{E}_N} u_N^k$ has the same regularity as u_N^k because it is equal to $u_N^k - \chi_{\mathcal{E}_N^k}$. We prove in Lemmata 9.1, 9.2 that u_N^k is close to $\chi_{\mathcal{E}_N^k}$ and that the local ergodic theorem holds for $\mathcal{L}_{\mathcal{E}_N} u_N^k$ if γ_N is larger than the equilibration times in the valleys and smaller than the transition times between valleys.

9.1. The enlarged process

We assume below that the reader is familiar with the results on enlarged and reflected chains summarized in Section A.

We do not require below the process $\eta_N(t)$ to be reversible, but we impose certain conditions on the reflected processes. Denote by $\eta_N^{R,k}(t)$ the process $\eta_N(t)$ reflected at $\mathcal{E}_N^k$. Recall that this means that we forbid all jumps between $\mathcal{E}_N^k$ and its complement, and consider the resulting dynamics in $\mathcal{E}_N^k$.

Denote by $\pi_{\mathcal{E}^k}$ the stationary measure π_N conditioned to $\mathcal{E}_N^k$. We assume that for all $k \in S$ the reflected process at $\mathcal{E}_N^k$ is *irreducible* and that $\pi_{\mathcal{E}^k}$ is a stationary state (and therefore the unique stationary state up to multiplicative constants). If the process is reversible, the second condition follows from the first one. By Lemma A.7, this is also the case in the non-reversible setting if the valley $\mathcal{E}_N^k$ is formed by cycles.

Denote by $\mathcal{L}_N^{R,k}$ the generator of the process $\eta_N^{R,k}(t)$ and by $t_{\mathrm{rel}}^{k,N}$ the *relaxation time* of the symmetric part of the generator:

$$\frac{1}{t_{\mathrm{rel}}^{k,N}} = \inf_f \frac{\langle (-\mathcal{L}_N^{R,k}) f , f \rangle_{\pi_{\mathcal{E}^k}}}{\langle f , f \rangle_{\pi_{\mathcal{E}^k}}} ,$$

where the infimum is carried over all zero-mean functions $f : \mathcal{E}_N^k \to \mathbb{R}$.

Let $\mathcal{E}_N^{\star,k}$ be copies of the sets $\mathcal{E}_N^k$, $k \in S$, and set

$$\mathcal{E}_N^\star \ := \ \bigcup_{k \in S} \mathcal{E}_N^{\star,k} \,, \quad \breve{\mathcal{E}}_N^{\star,k} \ := \ \bigcup_{j \neq k} \mathcal{E}_N^{\star,j} \,.$$

Denote by $P_\star : \mathcal{E}_N \cup \mathcal{E}_N^\star \to \mathcal{E}_N \cup \mathcal{E}_N^\star$ the application which maps a configuration in $\mathcal{E}_N$, $\mathcal{E}_N^\star$, to its copy in $\mathcal{E}_N^\star$, $\mathcal{E}_N$, respectively.

Fix a sequence γ_N, and denote by $\zeta_N(t)$ the γ_N-*enlargement* of the trace process $\eta^{\mathcal{E}_N}(t)$. The process $\zeta_N(t)$ is a Markov chain taking values in $\mathcal{E}_N \cup \mathcal{E}_N^\star$ and whose generator, denoted by $\mathcal{L}_{\mathcal{E}_N,\star}$, is given by

$$(\mathcal{L}_{\mathcal{E}_N,\star} f)(\eta) = \sum_{\xi \in \mathcal{E}_N} R_N^T(\eta,\xi) \left\{ f(\xi) - f(\eta) \right\} + \frac{1}{\gamma_N} \left\{ f(P_\star \eta) - f(\eta) \right\}, \quad \eta \in \mathcal{E}_N,$$

$$(\mathcal{L}_{\mathcal{E}_N,\star} f)(\eta) = \frac{1}{\gamma_N} \left\{ f(P_\star \eta) - f(\eta) \right\}, \quad \eta \in \mathcal{E}_N^\star.$$

In this formula, $R_N^T(\eta,\xi)$ represents the jump rates of the trace process $\eta^{\mathcal{E}_N}(t)$. Hence, from a configuration $\eta \in \mathcal{E}_N^\star$ the chain may only jump to $P_\star \eta$ and this happens at rate $1/\gamma_N$. From a configuration $\eta \in \mathcal{E}_N$, besides the jumps of the original chain, the enlarged process may also jump to $P_\star \eta$ and this happens at rate $1/\gamma_N$. The parameter γ_N will be large, which makes the jumps between $\mathcal{E}_N^\star$ and $\mathcal{E}_N$ rare.

The stationary state of $\zeta_N(t)$, denoted by $\pi_{\mathcal{E}}^\star$, is given by

$$\pi_{\mathcal{E}}^\star(\eta) \ = \ \pi_{\mathcal{E}}^\star(P_\star \eta) \ = \ (1/2)\,\pi_{\mathcal{E}}(\eta)\,, \quad \eta \in \mathcal{E}_N\,,$$

where, recall, $\pi_{\mathcal{E}}$ stands for the stationary state π_N conditioned to $\mathcal{E}_N$.

In dynamics in which the process jumps to a new valley before visiting all configurations in the valley, as configurations are not visited, it makes more sense to suppose that the dynamics starts from a distribution rather than from a configuration. Denote this initial distribution by ν_N and assume that there exist $\ell \in S$ and a finite constant C_0 such that for all $N \geq 1$

$$\nu_N(\mathcal{E}_N^\ell) \ = \ 1 \quad \text{and} \quad E_{\pi_{\mathcal{E}}}\left[\left(\frac{d\nu_N}{d\pi_{\mathcal{E}}} \right)^2 \right] \ \leq \ \frac{C_0}{\pi_{\mathcal{E}}(\mathcal{E}_N^\ell)} \,. \tag{9.2}$$

Note that the measure $\pi_{\mathcal{E}^\ell}$ satisfies this condition.

For two non-empty, disjoint subsets $\mathcal{A}$, $\mathcal{B}$ of $\mathcal{E}_N \cup \mathcal{E}_N^*$, denote by $\mathrm{cap}_\star(\mathcal{A}, \mathcal{B})$ the capacity between $\mathcal{A}$ and $\mathcal{B}$ for the enlarged process. Consider two sequences $(a_N : N \geq 1)$, $(b_N : N \geq 1)$ of positive real numbers. We say that a_N is much smaller than b_N, $a_N \ll b_N$, if $\lim_{N \to \infty} a_N / b_N = 0$.

Lemma 9.1. *Fix $k \in S$, two sequences of positive numbers γ_N, θ_N, and a sequence of probability measures ν_N satisfying* (9.2). *Assume that $\gamma_N \ll \theta_N$ and that there exists a finite constant C_0 such that for all $N \geq 1$*

$$\frac{\mathrm{cap}_\star(\mathcal{E}_N^{\star,k}, \breve{\mathcal{E}}_N^{\star,k})}{\pi_{\mathcal{E}}^\star(\mathcal{E}_N^{\star,k})} \ \leq \ \frac{C_0}{\theta_N} \,. \tag{9.3}$$

Then, representing by u_N^k the solution of the resolvent equation (9.1),

$$\lim_{N\to\infty} \sup_{t\geq 0} \mathbb{E}_{\nu_N}\Big[\,\big|\,\chi_{\mathcal{E}_N^k}(\eta^{\mathcal{E}}(t\theta_N)) - u_N^k(\eta^{\mathcal{E}}(t\theta_N))\,\big|\,\Big] \;=\; 0\;.$$

Denote by $\zeta^{\mathcal{E}_N^\star}(t)$ the trace of the process $\zeta_N(t)$ on $\mathcal{E}_N^\star$, and by $\mathbb{P}_\eta^{\star,\gamma_N}$, $\eta \in \mathcal{E}_N \cup \mathcal{E}_N^\star$, the probability measure on $D([0,\infty), \mathcal{E}_N \cup \mathcal{E}_N^\star)$ induced by the enlarged process starting from η. Let $r_N^\star(j,k)$, $j \neq k \in S$, be the coarse-grained jump rates at which the trace process $\zeta^{\mathcal{E}_N^\star}(t)$ jumps from $\mathcal{E}_N^{\star,j}$ to $\mathcal{E}_N^{\star,k}$. By (5.5), these rates are given by

$$r_N^\star(j,k) \;=\; \frac{1}{\pi_{\mathcal{E}}^\star(\mathcal{E}_N^{\star,j})} \sum_{\eta \in \mathcal{E}_N^{\star,j}} \pi_{\mathcal{E}}^\star(\eta)\,\lambda_\star(\eta)\,\mathbb{P}_\eta^{\star,\gamma_N}\big[\,H(\mathcal{E}_N^{\star,k}) < H^+(\check{\mathcal{E}}_N^{\star,j})\,\big]\;, \quad (9.4)$$

where $\lambda_\star(\eta)$ represents the holding rates of $\zeta_N(t)$. Since the enlarged process jumps from $\eta \in \mathcal{E}_N^\star$ to $P_\star\eta$ at rate γ_N, the previous expression is equal to

$$\frac{1}{\gamma_N\,\pi_{\mathcal{E}}(\mathcal{E}_N^j)} \sum_{\eta \in \mathcal{E}_N^j} \pi_{\mathcal{E}}(\eta)\,\mathbb{P}_\eta^{\star,\gamma_N}\big[\,H(\mathcal{E}_N^{\star,k}) < H(\check{\mathcal{E}}_N^{\star,j})\,\big]\;.$$

According to Section 6, in the reversible case, the coarse-grained jump rates $r_N^\star(j,k)$ can be expressed in terms of capacities, while in the non-reversible case they can be computed if there are good approximations of the equilibrium potential. Assume, from now on, that these rates converge: There exist a time-scale θ_N and jump rates $\boldsymbol{r}(j,k)$ such that

$$\lim_{N\to\infty} \theta_N\, r_N^\star(j,k) \;=\; \boldsymbol{r}(j,k) \quad \text{for all } j \neq k \in S\;. \qquad (9.5)$$

Condition (9.3) follows from this hypothesis since

$$\mathrm{cap}_\star(\mathcal{E}_N^{\star,k}, \check{\mathcal{E}}_N^{\star,k}) \;=\; \pi_{\mathcal{E}}^\star(\mathcal{E}_N^{\star,k}) \sum_{\ell \neq k} r_N^\star(k,\ell)\;.$$

The sequence θ_N represents the time-scale at which the process jumps between valleys. The proof of a metastable behavior is set up on the ground that this time-scale is much larger than the equilibration time inside the valleys. This hypothesis is formulated here by requiring the relaxation times of the processes reflected at a valleys to be much smaller than θ_N: for all $k \in S$,

$$t_{\mathrm{rel}}^{k,N} \;\ll\; \theta_N\;. \qquad (9.6)$$

Let $w_N^k = \mathcal{L}_{\mathcal{E}_N} u_N^k$. Recall from (5.3) that $\mathcal{G}_N$ represents the σ-algebra of subsets of $\mathcal{E}_N$ generated by the sets $\mathcal{E}_N^j$, $j \in S$. Let

$$\widehat{w}_N^k(\eta) \;=\; E_{\pi_{\mathcal{E}}}\big[\,w_N^k \,|\, \mathcal{G}_N\,\big]\;. \qquad (9.7)$$

Lemma 9.2. *Fix $\ell \in S$, and a sequence of probability measures ν_N satisfying* (9.2). *Assume that for all j, $k \in S$,*

$$\frac{\pi_N(\mathcal{E}_N^k)}{\pi_N(\mathcal{E}_N^\ell)} \, t_{\text{rel}}^{j,N} \;\ll\; \theta_N \;. \tag{9.8}$$

Let γ_N be a sequence such that $\max_{j,k} \alpha_N^{j,k} \ll \gamma_N \ll \theta_N$, where $\alpha_N^{j,k}$ stands for the left-hand side of (9.8). *Then, for all $T > 0$, $k \in S$,*

$$\lim_{N \to \infty} \mathbb{E}_{\nu_N} \Big[\sup_{t \le T\theta_N} \Big| \int_0^t \Big\{ w_N^k(\eta^{\mathcal{E}_N}(s)) - \widehat{w}_N^k(\eta^{\mathcal{E}_N}(s)) \Big\} ds \Big| \Big] \;=\; 0 \;.$$

By (9.12) and a straightforward computation,

$$\widehat{w}_N^k(\eta) \;=\; \sum_{j \in S} \boldsymbol{r}_N^\star(j,k) \, \chi_{\mathcal{E}_N^j}(\eta) \;,$$

where $\boldsymbol{r}_N^\star(j,k)$, $j \ne k$, are the coarse-grained jump rates introduced in (9.4), and $\boldsymbol{r}_N^\star(j,j) = - \sum_{k \ne j} \boldsymbol{r}_N^\star(j,k)$. Thus, for every function $F : S \to \mathbb{R}$,

$$\sum_{k \in S} F(k) \, \widehat{w}_N^k(\eta) \;=\; \sum_{j \in S} \chi_{\mathcal{E}_N^j}(\eta) \sum_{k \ne j} \boldsymbol{r}_N^\star(j,k) \left[F(k) - F(j) \right] . \tag{9.9}$$

Fix $\ell \in S$ and a sequence of probability measures satisfying conditions (9.2). Let $\mathbf{P}^N$ be the probability measure on $D([0,\infty), S)$ induced by the process $\boldsymbol{X}_N^T$ and the measure ν_N. Next theorem is the main result of this section.

Theorem 9.3. *Fix $\ell \in S$ and a sequence ν_N of probability measures satisfying* (9.2). *Assume that conditions* (9.5) *and* (9.8) *are in force. Then, every limit point $\mathbf{P}$ of the sequence $\mathbf{P}^N$ such that*

$$\mathbf{P}\big[\, X(t-) \,=\, X(t) \,\big] \quad \text{for all } t > 0 \;. \tag{9.10}$$

solves the (L, δ_ℓ) martingale problem, where L is the generator of the S-valued Markov chain whose jump rates are $\boldsymbol{r}(j,k)$.

Theorem 9.3 describes the asymptotic evolution of the trace of the Markov $\eta(t)$ on $\mathcal{E}_N$. The next lemma shows that in the time scale θ_N the time spent on the complement of $\mathcal{E}_N$ is negligible. The proof is similar to the one of Lemma 7.3 and uses Schwarz inequality and assumption (9.2) to replace ν_N by π_N.

Lemma 9.4. *Assume that*

$$\lim_{N \to \infty} \frac{\pi(\Delta_N)}{\pi(\mathcal{E}_N^j)} \;=\; 0$$

for all $j \in S$. Fix $\ell \in S$, and let $\{\nu_N : N \ge 1\}$ be a sequence of probability measures satisfying (9.2). *Then, for every $t > 0$,*

$$\lim_{N \to \infty} \mathbb{E}_{\nu_N} \Big[\int_0^t \chi_{\Delta_N}(\eta(s\theta_N)) \, ds \Big] \;=\; 0 \;.$$

Remark 9.5. *The introduction of the enlarged process is inspired by the definition of the* soft hitting time *of Bianchi and Gaudillière [24].*

Remark 9.6. *Hypothesis* (9.8) *can be divided in two: Assume* (9.6), *and suppose that there exist constants* $0 < c_0 < C_0 < \infty$ *such that* $c_0\,\pi_N(\mathcal{E}_N^k) \leq \pi_N(\mathcal{E}_N^\ell) \leq C_0\,\pi_N(\mathcal{E}_N^k)$ *for all* k, $\ell \in S$.

Remark 9.7. *Hypothesis* (9.8) *can be weaken as follows. Instead of fixing the same rate* γ_N *for all valleys, we may choose a valley-dependent rate. This does not alter the stationary state, and it permits to choose larger parameters* γ_N *for deeper valleys. Assumption* (9.8) *may also be weaken to admit a deep valley, all the other ones being shallow (cf. [20]).*

Remark 9.8. *In Subsection* 13.5, *we apply the method presented above to a polymer model examined by Caputo et al. in [37, 35]. It can also be employed to derive the reduced model of the random walk presented in Section* 2. *We refer to in [20]. Lacoin and Teixeira [84] followed this scheme to prove the metastable behavior of a polymer interface which interacts with an attractive substrate.*

Proof of Theorem 9.3. Fix $\ell \in S$ and a sequence ν_N of probability measures satisfying (9.2). Fix a function $F : S \to \mathbb{R}$ and a limit point **P** of the sequence $\mathbf{P}^N$ satisfying (9.10). Assume, without loss of generality, that $\mathbf{P}^N$ converges to **P**. We claim that

$$M^F(t) \; := \; F(X_t) \; - \; F(X_0) \; - \; \int_0^t (L\,F)(X_s)\,ds \tag{9.11}$$

is a martingale under **P**, where L is the generator associated to the jump rates $\boldsymbol{r}(j,k)$ introduced in (9.5).

Fix $0 \leq s < t$, $q \geq 1$, $0 \leq t_1 < \cdots < t_q \leq s$, and a bounded function $g : S^q \to \mathbb{R}$. Let $G = g(X(t_1),\ldots,X(t_q))$, where $X(s)$ represents the coordinate process of $D([0,\infty),S)$. We shall prove that

$$\mathbf{E}\left[\,M^F(t)\,G\,\right] \;=\; \mathbf{E}\left[\,M^F(s)\,G\,\right]\,,$$

where **E** stands for the expectation with respect to **P**.

Fix a sequence γ_N such that for all j, $k \in S$,

$$\frac{\pi_N(\mathcal{E}_N^k)}{\pi_N(\mathcal{E}_N^\ell)}\,t_{\mathrm{rel}}^{j,N} \;\ll\; \gamma_N \;\ll\; \theta_N\,,$$

which is possible in view of (9.8), and recall that we denote by u_N^k the solution of (9.1). Let

$$H_N(\eta) \;=\; \sum_{k\in S} F(k)\,u_N^k(\eta)\,,\quad \eta \in \mathcal{E}_N\,.$$

By the Markov property of the trace process $\eta^{\mathcal{E}}(t)$,

$$M_t^N \;=\; H_N(\eta^{\mathcal{E}}(t\theta_N)) \;-\; H_N(\eta^{\mathcal{E}}(0)) \;-\; \int_0^{t\theta_N} (\mathcal{L}_{\mathcal{E}_N}H_N)(\eta^{\mathcal{E}}(s))\,ds$$

is a martingale. In particular, if

$$G_N \; = \; g\big(\boldsymbol{X}_N^T(t_1) \,, \, \ldots \,, \, \boldsymbol{X}_N^T(t_q) \big) \; = \; g\big(X_N^T(t_1\theta_N) \,, \, \ldots \,, \, X_N^T(t_q\theta_N) \big) \,,$$

we have that

$$\mathbb{E}_{\nu_N} \big[\, M_t^N \, G_N \,\big] \; = \; \mathbb{E}_{\nu_N} \big[\, M_s^N \, G_N \,\big] \,,$$

so that

$$\mathbb{E}_{\nu_N} \Big[G_N \Big\{ H_N(\eta^{\mathcal{E}}(t\theta_N)) \, - \, H_N(\eta^{\mathcal{E}}(s\theta_N)) \, - \, \int_{s\theta_N}^{t\theta_N} (\mathcal{L}_{\mathcal{E}_N} H_N)(\eta^{\mathcal{E}}(r)) \, dr \Big\} \Big] \; = \; 0 \,.$$

By Lemma 9.1,

$$\lim_{N\to\infty} \sup_{r\geq 0} \mathbb{E}_{\nu_N} \Big[\, \big| \, H_N(\eta^{\mathcal{E}}(r\theta_N)) \, - \, (F \circ \Psi_N)(\eta^{\mathcal{E}}(r\theta_N)) \, \big| \, \Big] \; = \; 0 \,.$$

Thus, by the penultimate equation and since $\boldsymbol{X}_N^T(t) = X_N^T(t\theta_N) = \Psi_N(\eta^{\mathcal{E}_N}(t\theta_N))$,

$$\lim_{N\to\infty} \mathbb{E}_{\nu_N} \Big[G_N \Big\{ F(\boldsymbol{X}_N^T(t)) \, - \, F(\boldsymbol{X}_N^T(s)) \, - \, \int_{s\theta_N}^{t\theta_N} (\mathcal{L}_{\mathcal{E}_N} H_N)(\eta^{\mathcal{E}}(r)) \, dr \Big\} \Big] \; = \; 0 \,.$$

By definition of H_N and w_N^k, introduced just above (9.7), $\mathcal{L}_{\mathcal{E}_N} H_N = \sum_k F(k) \, w_N^k$. Hence, by Lemma 9.2 and (9.9),

$$\lim_{N\to\infty} \mathbb{E}_{\nu_N} \Big[G_N \Big\{ F(\boldsymbol{X}_N^T(t)) \, - \, F(\boldsymbol{X}_N^T(s)) \, - \, \int_s^t (L_N^\star F)(\boldsymbol{X}_N^T(r)) \, dr \Big\} \Big] \; = \; 0 \,.$$

where $L_N^\star$ is the generator of a S-valued Markov chain given by

$$(L_N^\star F)(j) \; = \; \sum_{k\in S} \theta_N \, r_N^\star(j,k) \, [\, F(k) - F(j) \,] \,.$$

At this point, the martingale has been expressed as a function of the process $\boldsymbol{X}_N^T(t)$. By definition of the measure $\mathbf{P}^N$, the previous expectation is equal to

$$\mathbf{E}^N \Big[g\big(X(t_1) \,, \, \ldots \,, \, X(t_q) \big) \Big\{ F(X(t)) \, - \, F(X(s)) \, - \, \int_s^t (L_N^\star F)(X(r)) \, dr \Big\} \Big] \,,$$

where, recall, $X(t)$ represents the coordinate process in $D([0,\infty), S)$ and $\mathbf{E}^N$ expectation with respect to $\mathbf{P}^N$.

By assumption (9.5), $(L_N^\star F)(k)$ converges to $(LF)(k)$ for all $k \in S$. Therefore, as $\mathbf{P}^N$ converges to $\mathbf{P}$ and in view of (9.10) [which guarantees that the finite-dimensional projections are continuous], passing to the limit, we get that

$$\mathbf{E} \Big[g\big(X(t_1) \,, \, \ldots \,, \, X(t_q) \big) \Big\{ F(X(t)) \, - \, F(X(s)) \, - \, \int_s^t (LF)(X(r)) \, dr \Big\} \Big] \; = \; 0 \,.$$

This shows that (9.11) holds, and completes the proof of the theorem. $\quad\square$

9.2. The resolvent equation

We examine in this subsection the asymptotic behavior of the solution of resolvent equation (9.1).

Fix $\gamma_N > 0$ and consider the γ_N-enlargement of the process $\eta^{\mathcal{E}_N}(t)$. Let $h_N^k : \mathcal{E}_N \cup \mathcal{E}_N^\star \to [0,1]$ be the equilibrium potential between the sets $\mathcal{E}_N^{\star,k}$ and $\breve{\mathcal{E}}_N^{\star,k}$:

$$h_N^k(\eta) \;:=\; \mathbb{P}_\eta^{\star,\gamma_N}\left[\, H(\mathcal{E}_N^{\star,k}) \,<\, H(\breve{\mathcal{E}}_N^{\star,k}) \,\right]. \tag{9.12}$$

Since $\mathcal{L}_{\mathcal{E}_N,\star} h_N^k = 0$ on $\mathcal{E}_N$, we deduce that the restriction of h_N^k to $\mathcal{E}_N$ solves the resolvent equation (9.1). Since the solution is unique, $u_N^k = h_N^k$ on $\mathcal{E}_N$ and we have a simple stochastic representation of the solution of the resolvent equations.

Remark 9.9. *The enlargement of the chain $\eta^{\mathcal{E}_N}(t)$ thus provides a stochastic representation of the resolvent equation* (9.1).

Lemma 9.10. *There exists a finite constant C_0, independent of N, such that for all $k \in S$,*

$$\frac{1}{\gamma_N} \sum_{\eta \in \mathcal{E}_N^k} \pi_{\mathcal{E}}(\eta)\,[1 - u_N^k(\eta)]^2 + \langle\, (-\mathcal{L}_{\mathcal{E}_N})\, u_N^k,\, u_N^k \,\rangle_{\pi_{\mathcal{E}}} + \frac{1}{\gamma_N} \sum_{\eta \in \breve{\mathcal{E}}_N^k} \pi_{\mathcal{E}}(\eta)\, u_N^k(\eta)^2$$
$$\leq \frac{C_0}{\theta_N}\, \pi_{\mathcal{E}}(\mathcal{E}_N^k).$$

Proof. Denote the left-hand side of the inequality by A_N, and by B_N the same expression with $\pi_{\mathcal{E}}$ in place of $\pi_{\mathcal{E}}^\star$. Since $\pi_{\mathcal{E}}^\star(\eta) = (1/2)\, \pi_{\mathcal{E}}(\eta)$, $\eta \in \mathcal{E}_N$, $A_N = 2B_N$. As u_N^k and h_N^k coincide on $\mathcal{E}_N$, we may replace the former by the latter. On the other hand, as $h_N^k = \chi_{\mathcal{E}_N^{k,\star}}$ on $\mathcal{E}_N^\star$, $B_N = D_{N,\star}(h_N^k)$, where $D_{N,\star}(f)$ represents the Dirichlet form of f with respect to the enlarged process $\zeta_N(t)$.

By (B.7),

$$D_{N,\star}(h_N^k) \;=\; \mathrm{cap}_\star(\mathcal{E}_N^{\star,k}, \breve{\mathcal{E}}_N^{\star,k})\,.$$

Thus, $A_N = 2\,\mathrm{cap}_\star(\mathcal{E}_N^{\star,k}, \breve{\mathcal{E}}_N^{\star,k})$. By assumption (9.3), the capacity is less than or equal to $C_0\, \pi_{\mathcal{E}}^\star(\mathcal{E}_N^{\star,k})/\theta_N$ for some finite constant C_0. This proves the assertion because $\pi_{\mathcal{E}}^\star(\mathcal{E}_N^{\star,k}) = (1/2)\, \pi_{\mathcal{E}}(\mathcal{E}_N^k)$. $\square$

Proof of Lemma 9.1. Fix $\ell \in S$, a sequence of probability measures ν_N satisfying the hypotheses of the lemma and $t > 0$. Denote by $S_{\mathcal{E}}(t)$, $t \geq 0$, the semigroup associated to the trace process $\eta^{\mathcal{E}}(t)$, and by f_t^N the Radon-Nikodym derivative $d\nu_N S_{\mathcal{E}}(t)/d\pi_{\mathcal{E}}$. By (A.9), $E_{\pi_{\mathcal{E}}}[(f_t^N)^2] \leq E_{\pi_{\mathcal{E}}}[(f_0^N)^2]$. Hence, by Schwarz inequality, the square of the expectation appearing in the statement of the lemma is bounded above by

$$E_{\pi_{\mathcal{E}}}\left[\left(\frac{d\nu_N}{d\pi_{\mathcal{E}}}\right)^2\right] E_{\pi_{\mathcal{E}}}\left[\left(\chi_{\mathcal{E}_N^\ell} - u_N^\ell\right)^2\right].$$

By Lemma 9.10, the second term is bounded by $C_0 \, \gamma_N \, \pi_{\mathcal{E}}(\mathcal{E}_N^\ell)/\theta_N$. Thus, by the assumption on the sequence of probability measures ν_N, the previous displayed formula is bounded by $C_0 \, \gamma_N/\theta_N$. This expression vanishes as $N \to \infty$ by the hypothesis on γ_N. $\qquad\square$

9.3. Local ergodicity

The proof of Lemma 9.2 is divided in several steps. Denote by $\langle \, \cdot \, , \, \cdot \, \rangle_{\pi_{\mathcal{E}}}$ the scalar product in $L^2(\pi_{\mathcal{E}})$. For a zero-mean function $f : \mathcal{E}_N \to \mathbb{R}$, let $\|f\|_{-1}$ be the $\mathcal{H}_{-1}$ norm of f associated to the generator $\mathcal{L}_{\mathcal{E}_N}$:

$$\|f\|_{-1}^2 \;=\; \sup_h \Big\{ 2 \, \langle \, f \, , \, h \, \rangle_{\pi_{\mathcal{E}}} \, - \, \langle \, h \, , \, (-\mathcal{L}_{\mathcal{E}_N}) \, h \, \rangle_{\pi_{\mathcal{E}}} \Big\} \,,$$

where the supremum is carried over all functions $h : \mathcal{E}_N \to \mathbb{R}$. By [82, Lemma 2.4], for every function $f : \mathcal{E}_N \to \mathbb{R}$ which has zero-mean with respect to $\pi_{\mathcal{E}}$, and every $T > 0$,

$$\mathbb{E}_{\pi_{\mathcal{E}}} \Big[\sup_{0 \le t \le T} \Big(\int_0^t f(\eta^{\mathcal{E}_N}(s)) \, ds \Big)^2 \Big] \;\le\; 24 \, T \, \|f\|_{-1}^2 \,. \tag{9.13}$$

Recall that we denote by $\pi_{\mathcal{E}^k}$ the stationary measure π_N conditioned to $\mathcal{E}_N^k$. Let $\mathcal{L}_{R,\mathcal{E}_N^k}$ be the generator of the reflected process $\eta_N(t)$ at $\mathcal{E}_N^k$. For a function $f : \mathcal{E}_N^k \to \mathbb{R}$ which has zero-mean with respect to $\pi_{\mathcal{E}^k}$, denote by $\|f\|_{k,-1}$ the $\mathcal{H}_{-1}$ norm of f with respect to the generator $\mathcal{L}_{R,\mathcal{E}_N^k}$:

$$\|f\|_{k,-1}^2 \;=\; \sup_h \Big\{ 2 \, \langle \, f \, , \, h \, \rangle_{\pi_{\mathcal{E}^k}} \, - \, \langle \, h \, , \, (-\mathcal{L}_{R,\mathcal{E}_N^k}) \, h \, \rangle_{\pi_{\mathcal{E}^k}} \Big\} \,,$$

where the supremum is carried over all functions $h : \mathcal{E}_N^k \to \mathbb{R}$. It is clear that

$$\sum_{j \in S} \pi_{\mathcal{E}}(\mathcal{E}_N^j) \, \langle \, h \, , \, (-\mathcal{L}_{R,\mathcal{E}_N^j}) \, h \, \rangle_{\pi_{\mathcal{E}^j}} \;\le\; \langle \, h \, , \, (-\mathcal{L}_{\mathcal{E}_N}) \, h \, \rangle_{\pi_{\mathcal{E}}}$$

for any function $h : \mathcal{E}_N \to \mathbb{R}$. These expression are not equal because two kinds of jumps appear on the right-hand side and do not on the left: The trace process may jump between valleys, and it may also perform a jump inside a valley (crossing the set Δ_N) which is not possible in the original dynamics.

It follows from the previous inequality and from the formulae for the $\mathcal{H}_{-1}$ norms that for every function $f : \mathcal{E}_N \to \mathbb{R}$ which has zero-mean with respect to each measure $\pi_{\mathcal{E}^j}$,

$$\|f\|_{-1}^2 \;\le\; \sum_{j \in S} \pi_{\mathcal{E}}(\mathcal{E}_N^j) \, \|f\|_{j,-1}^2 \,. \tag{9.14}$$

Lemma 9.11. *Let* $\{\nu_N : N \ge 1\}$ *be a sequence of probability measures on* $\mathcal{E}_N$. *Then, for every function* $f : \mathcal{E}_N \to \mathbb{R}$ *which has zero-mean with respect to each measure* $\pi_{\mathcal{E}^j}$ *and for every* $T > 0$,

$$\Big(\mathbb{E}_{\nu_N} \Big[\sup_{t \le T} \Big| \int_0^t f(\eta^{\mathcal{E}}(s)) \, ds \Big| \Big] \Big)^2 \;\le\; 24 \, T \, E_{\pi_{\mathcal{E}}} \Big[\Big(\frac{\nu_N}{\pi_{\mathcal{E}}} \Big)^2 \Big] \sum_{j \in S} \pi_{\mathcal{E}}(\mathcal{E}_N^j) \, \|f\|_{j,-1}^2 \,.$$

Proof. By Schwarz inequality, the expression on the left hand side is bounded above by

$$E_{\pi_{\mathcal{E}}} \Big[\Big(\frac{\nu_N}{\pi_{\mathcal{E}}} \Big)^2 \Big] \, \mathbb{E}_{\pi_{\mathcal{E}}} \Big[\sup_{t \le T} \Big(\int_0^t f(\eta^{\mathcal{E}}(s)) \, ds \Big)^2 \Big] \, .$$

By (9.13) and by (9.14), the second expectation is bounded by

$$24 \, T \sum_{j \in S} \pi_{\mathcal{E}}(\mathcal{E}_N^j) \, \|f\|_{j,-1}^2 \, ,$$

as claimed. $\square$

Proof of Lemma 9.2. Fix $\ell \in S$, and a sequence of probability measures ν_N satisfying the hypotheses of the lemma. Fix $k \in S$. Since $w_N^k - \widehat{w}_N^k$ has zero-mean with respect to each $\pi_{\mathcal{E}^j}$, by the assumption on the sequence ν_N and Lemma 9.11, the square of the expectation appearing in the statement of the lemma is bounded by

$$\frac{C_0 \, T \, \theta_N}{\pi_{\mathcal{E}}(\mathcal{E}^\ell)} \sum_{j \in S} \pi_{\mathcal{E}}(\mathcal{E}_N^j) \, \|w_N^k - \widehat{w}_N^k\|_{j,-1}^2 \tag{9.15}$$

for some finite constant C_0.

By (9.1), on the set $\mathcal{E}_N^k$, $\mathcal{L}_{\mathcal{E}_N} u_N^k = -(1/\gamma_N)(1 - u_N^k)$, so that $w_N^k - \widehat{w}_N^k = (1/\gamma_N)(u_N^k - \widehat{u}_N^k)$. Hence, by the spectral gap of the reflected process,

$$\| \, w_N^k - \widehat{w}_N^k \, \|_{k,-1}^2 \; = \; \frac{1}{\gamma_N^2} \, \| \, u_N^k - \widehat{u}_N^k \, \|_{k,-1}^2 \; \le \; \frac{t_{\mathrm{rel}}^{k,N}}{\gamma_N^2} \, \| \, u_N^k - \widehat{u}_N^k \, \|_{\pi_{\mathcal{E}^k}}^2 \, .$$

Since $\| \, u_N^k - \widehat{u}_N^k \, \|_{\pi_{\mathcal{E}^k}}^2 \le \| \, u_N^k - 1 \, \|_{\pi_{\mathcal{E}^k}}^2$, by Lemma 9.10,

$$\| \, w_N^k - \widehat{w}_N^k \, \|_{k,-1}^2 \; \le \; C_0 \, \frac{t_{\mathrm{rel}}^{k,N}}{\gamma_N \, \theta_N}$$

for some finite constant C_0.

Similarly, since $\mathcal{L}_{\mathcal{E}_N} u_N^k = (1/\gamma_N) \, u_N^k$ on the sets $\mathcal{E}_N^j$, $j \ne k$,

$$\| \, w_N^k - \widehat{w}_N^k \, \|_{j,-1}^2 \; = \; \frac{1}{\gamma_N^2} \, \| \, u_N^k - \widehat{u}_N^k \, \|_{j,-1}^2 \; \le \; C_0 \, \frac{\pi_{\mathcal{E}}(\mathcal{E}_N^k)}{\pi_{\mathcal{E}}(\mathcal{E}_N^j)} \, \frac{t_{\mathrm{rel}}^{j,N}}{\gamma_N \, \theta_N} \, .$$

Therefore, the sum appearing in (9.15) is bounded by

$$C_0 \, T \, |S| \, \frac{\max_{j \in S} t_{\mathrm{rel}}^{j,N}}{\gamma_N} \, \frac{\pi_N(\mathcal{E}_N^k)}{\pi_N(\mathcal{E}_N^\ell)} \, .$$

By the hypotheses of the lemma, this expression vanishes as $N \uparrow \infty$, which completes the proof. $\square$

Proof of Lemma 9.4.. Fix $\ell \in S$, and let ν_N be a sequence of probability measures satisfying (9.2). By Schwarz inequality, the square of the expectation appearing in the statement of the lemma is bounded above by

$$\frac{1}{\pi_N(\mathcal{E}_N)} \, E_{\pi_\mathcal{E}} \Big[\Big(\frac{d\nu_N}{d\pi_\mathcal{E}} \Big)^2 \Big] \, \mathbb{E}_{\pi_N} \Big[\Big(\int_0^t \chi_{\Delta_N} \big(\eta(s\theta_N) \big) \, ds \Big)^2 \Big]$$

By assumption (9.2), the first expectation is bounded by $C_0/\pi_\mathcal{E}(\mathcal{E}_N^\ell)$. On the other hand, by Schwarz inequality, the second expectation is less than or equal to

$$t \, \mathbb{E}_{\pi_N} \Big[\int_0^t \chi_{\Delta_N} \big(\eta(s\theta_N) \big) \, ds \Big] \;=\; t^2 \pi_N(\Delta_N) \,.$$

The expression appearing in the penultimate displayed formula is thus bounded above by $C_0 \, t^2 \, [\, \pi_N(\Delta_N)/\pi_N(\mathcal{E}_N^\ell) \,]$, which concludes the proof of the lemma. $\square$

10. Tightness

In this section, we present sufficient conditions for the tightness of the sequence $\mathbf{P}^N$ introduced in Theorems 4.2, 8.2 and 9.3. We need a slight generalization of Lemma 8.1. Recall the notation introduced just before this lemma. We proved there that for each $t \geq 0$ and $\eta \in E_N$, $S_{\mathcal{E}_N}(t)$ is a stopping time with respect to the filtration $(\mathcal{F}_t^\eta : t \geq 0)$.

Lemma 10.1. *Let $\{\mathcal{G}_r : r \geq 0\}$ be the filtration given by $\mathcal{G}_r = \mathcal{F}_{S_{\mathcal{E}_N}(r)}^\eta$, and let τ be a stopping time with respect to $\{\mathcal{G}_r\}$. Then, $S_{\mathcal{E}_N}(\tau)$ is a stopping time with respect to $\{\mathcal{F}_t^\eta\}$.*

Proof. Fix a stopping time τ with respect to the filtration $\{\mathcal{G}_r\}$. This means that for every $t \geq 0$, $\{\tau \leq t\} \in \mathcal{G}_t = \mathcal{F}_{S_{\mathcal{E}_N}(t)}^\eta$. Hence, for all $r \geq 0$,

$$\{\tau \leq t\} \cap \{S_{\mathcal{E}_N}(t) \leq r\} \, \in \, \mathcal{F}_r^\eta \,.$$

We claim that $\{S_{\mathcal{E}_N}(\tau) < t\} \in \mathcal{F}_t^\eta$. Indeed, by (8.3), this event is equal to $\{T_{\mathcal{E}_N}(t) > \tau\}$, which can be written as

$$\bigcup_{q \in \mathbb{Q}} \{\tau \leq q\} \cap \{T_{\mathcal{E}_N}(t) > q\} \;=\; \bigcup_{q \in \mathbb{Q}} \{\tau \leq q\} \cap \{S_{\mathcal{E}_N}(q) < t\}$$
$$=\; \bigcup_{q \in \mathbb{Q}} \bigcup_{n \geq 1} \{\tau \leq q\} \cap \{S_{\mathcal{E}_N}(q) \leq t - (1/n)\} \,.$$

By the penultimate displayed equation, each term belongs to $\mathcal{F}_{t-(1/n)}^\eta \subset \mathcal{F}_t^\eta$, which proves the claim.

We may conclude. Since

$$\{S_{\mathcal{E}_N}(\tau) \leq t\} \;=\; \bigcap_q \{S_{\mathcal{E}_N}(\tau) < t + q\} \,,$$

where the intersection is carried out over all $q \in (0, \infty) \cap \mathbb{Q}$, and since the filtration $\{\mathcal{F}_t^\eta\}$ is right continuous, by the previous claim, $\{S_{\mathcal{E}_N}(\tau) \leq t\} \in \mathcal{F}_t^\eta$. $\qquad\square$

Recall that $\xi_N(t) = \eta_N(t\theta_N)$, and the definition of the measure $\mathbb{Q}_\eta^N$ introduced just before Lemma 8.1. Expectation with respect to this measure is denoted by $\mathbb{Q}_\eta^N$, as well. Note that $\eta^{\mathcal{E}_N}(t\theta_N) = \xi^{\mathcal{E}_N}(t)$.

Lemma 10.2. *Suppose that for all $t > 0$,*

$$\lim_{N\to\infty} \max_{\xi\in\mathcal{E}_N} \mathbb{Q}_\xi^N \Big[\int_0^t \chi_{\Delta_N}(\xi_N(s))\, ds \Big] = 0 \,, \tag{10.1}$$

and that

$$\lim_{\delta\to 0} \limsup_{N\to\infty} \max_{j\in S} \max_{\xi\in\mathcal{E}_j^N} \mathbb{Q}_\xi^N \big[H(\breve{\mathcal{E}}_N^j) \leq \delta \big] = 0 \,. \tag{10.2}$$

Then, the sequence of measures $\mathbf{P}^N$ is tight. Moreover, every limit point $\mathbf{P}$ is such that

$$\mathbf{P}\big[X(t) \neq X(t-) \big] = 0$$

for every $t > 0$.

Proof. Fix $\eta \in \mathcal{E}_N$. According to Aldous' criterion [26], we have to show that for every $\delta > 0$, $R > 0$,

$$\lim_{a_0\to 0} \limsup_{\epsilon\to 0} \sup \mathbb{P}_\eta^N \big[\big| \boldsymbol{X}_N^T(\tau + a) - \boldsymbol{X}_N^T(\tau) \big| > \delta \big] = 0 \,,$$

where the supremum is carried over all stopping times τ bounded by R and all $0 \leq a < a_0$. Since $\boldsymbol{X}_N^T(t) = \Psi_N(\xi^{\mathcal{E}_N}(t))$, the previous probability can be written as

$$\mathbb{Q}_\eta^N \big[\big| \Psi_N\big(\xi^{\mathcal{E}_N}(\tau + a)\big) - \Psi_N\big(\xi^{\mathcal{E}_N}(\tau)\big) \big| > \delta \big] \,.$$

Since $|\Psi_N\big(\xi^{\mathcal{E}_N}(\tau + a)\big) - \Psi_N\big(\xi^{\mathcal{E}_N}(\tau)\big)| > \delta$ entails that $\Psi_N\big(\xi^{\mathcal{E}_N}(\tau + a)\big) \neq \Psi_N\big(\xi^{\mathcal{E}_N}(\tau)\big)$, the expression in the previous displayed equation is bounded by

$$\mathbb{Q}_\eta^N \big[\Psi_N\big(\xi^{\mathcal{E}_N}(\tau + a)\big) \neq \Psi_N\big(\xi^{\mathcal{E}_N}(\tau)\big) \big] \,.$$

Fix $b = 2a_0$ so that $b - a \geq a_0$. Decompose this probability according to the event $\{S_{\mathcal{E}_N}(\tau + a) - S_{\mathcal{E}_N}(\tau) > b\}$ and its complement.

Suppose that $S_{\mathcal{E}_N}(\tau+a) - S_{\mathcal{E}_N}(\tau) > b$. In this case, $S_{\mathcal{E}_N}(\tau) + b < S_{\mathcal{E}_N}(\tau+a)$, so that $T_{\mathcal{E}_N}(S_{\mathcal{E}_N}(\tau)+b) \leq T_{\mathcal{E}_N}(S_{\mathcal{E}_N}(\tau+a)) = \tau+a$. Hence, as $T_{\mathcal{E}_N}(S_{\mathcal{E}_N}(t)) = t$, $T_{\mathcal{E}_N}(S_{\mathcal{E}_N}(\tau) + b) - T_{\mathcal{E}_N}(S_{\mathcal{E}_N}(\tau)) \leq a$, that is,

$$\int_{S_{\mathcal{E}_N}(\tau)}^{S_{\mathcal{E}_N}(\tau)+b} \chi_{\mathcal{E}_N}(\xi_N(s))\, ds \leq a \,.$$

In other words,

$$\int_{S_{\mathcal{E}_N}(\tau)}^{S_{\mathcal{E}_N}(\tau)+b} \chi_{\Delta_N}(\xi_N(s))\, ds \geq b - a \,.$$

By Lemma 10.2, $S_\mathcal{E}(\tau)$ is a stopping time for the filtration $\{\mathcal{F}_t^\eta\}$. Hence, by the strong Markov property and since $\xi_N(S_{\mathcal{E}_N}(t))$ belongs to $\mathcal{E}_N$ for all $t \geq 0$,

$$\mathbb{Q}_\eta^N\big[\, S_{\mathcal{E}_N}(\tau+a) - S_{\mathcal{E}_N}(\tau) > b \,\big] \;\leq\; \mathbb{Q}_\eta^N\Big[\, \int_{S_{\mathcal{E}_N}(\tau)}^{S_{\mathcal{E}_N}(\tau)+b} \chi_{\Delta_N}(\xi_N(s))\, ds \;\geq\; b-a \,\Big]$$

$$\leq\; \max_{\xi\in\mathcal{E}_N} \mathbb{Q}_\xi^N\Big[\, \int_0^b \chi_{\Delta_N}(\xi_N(s))\, ds \;\geq\; b-a \,\Big].$$

By Chebychev inequality, a change of variables and by our choice of b, this expression is less than or equal to

$$\frac{1}{(b-a)} \max_{\xi\in\mathcal{E}_N} \mathbb{Q}_\xi^N\Big[\, \int_0^b \chi_{\Delta_N}(\xi_N(s))\, ds \,\Big] \;\leq\; \frac{1}{a_0} \max_{\xi\in\mathcal{E}_N} \mathbb{Q}_\xi^N\Big[\, \int_0^{2a_0} \chi_{\Delta_N}(\xi_N(s))\, ds \,\Big].$$

By assumption (10.1), this expression vanishes as $N \to \infty$ for every $a_0 > 0$.

We turn to the case $\{S_{\mathcal{E}_N}(\tau+a) - S_{\mathcal{E}_N}(\tau) \leq b\}$. On this set we have that

$$\big\{\, \Psi_N(\xi_N(S_{\mathcal{E}_N}(\tau+a))) \neq \Psi_N(\xi_N(S_{\mathcal{E}_N}(\tau))) \,\big\}$$
$$\subset\; \big\{\, \Psi_N(\xi_N(S_{\mathcal{E}_N}(\tau)+c)) \neq \Psi_N(X(S_\mathcal{E}(\tau))) \text{ for some } 0 \leq c \leq b \,\big\}.$$

Since $S_{\mathcal{E}_N}(\tau)$ is a stopping time for the filtration $\{\mathcal{F}_t\}$ and since $\xi_N(S_{\mathcal{E}_N}(t))$ belongs to $\mathcal{E}_N$ for all t,

$$\mathbb{Q}_\eta^N\Big[\Psi_N(\xi_N(S_{\mathcal{E}_N}(\tau+a))) \neq \Psi(\xi_N(S_{\mathcal{E}_N}(\tau))),\, S_{\mathcal{E}_N}(\tau+a) - S_\mathcal{E}(\tau) \leq b\Big]$$
$$\leq\; \max_{\xi\in\mathcal{E}_N} \mathbb{Q}_\xi^N\Big[\Psi_N(\xi_N(c)) \neq \Psi_N(\xi) \text{ for some } 0 \leq c \leq b\Big].$$

If $\xi \in \mathcal{E}_N^j$, this later event corresponds to the event $\{H(\breve{\mathcal{E}}_N^j) \leq b\}$. The maximum is thus bounded by

$$\max_{j\in S} \max_{\xi\in\mathcal{E}_N^j} \mathbb{Q}_\xi^N\big[\, H(\breve{\mathcal{E}}_N^j) \leq b \,\big] \;=\; \max_{j\in S} \sup_{\xi\in\mathcal{E}_N^j} \mathbb{Q}_\xi^N\big[\, H(\breve{\mathcal{E}}_N^j) \leq 2a_0 \,\big].$$

By assumption (10.2), this expression vanishes as $N \to \infty$ and then $a_0 \to 0$. This completes the proof of the tightness.

The same argument shows that for every $t > 0$,

$$\lim_{a_0\to 0} \limsup_{N\to\infty} \mathbf{P}^N\big[\, X(t-a) \neq X(t) \text{ for some } 0 \leq a \leq a_0 \,\big] \;=\; 0.$$

Hence, if $\mathbf{P}$ is a limit point of the sequence $\mathbf{P}^N$,

$$\lim_{a_0\to 0} \mathbf{P}\big[\, X(t-a) \neq X(t) \text{ for some } 0 \leq a \leq a_0 \,\big] \;=\; 0.$$

This completes the proof of the second assertion of the lemma since $\{X(t) \neq X(t-)\} \subset \{X(t-a) \neq X(t) \text{ for some } 0 \leq a \leq a_0\}$ for all $a_0 > 0$. $\qquad\square$

Conditions (10.1), (10.2), can be formulated in terms of capacities. Next results is Theorem 2.6 in [15] and Theorem 2.1 in [19]. Note that we do not require the process to be reversible.

Theorem 10.3. *Assume that condition* (5.4) *is in force: For all* $j \in S$, *there exists* $\xi^{j,N} \in \mathcal{E}_N^j$ *such that*

$$\lim_{N \to \infty} \max_{\eta \in \mathcal{E}_N^j, \, \eta \neq \xi^{j,N}} \frac{\mathrm{cap}_N(\mathcal{E}_N^j, \breve{\mathcal{E}}_N^j)}{\mathrm{cap}_N(\eta, \xi^{j,N})} = 0 .$$

Assume, furthermore, that the coarse-grained jump rates converge: For all $j \neq k \in S$, *there exists* $\boldsymbol{r}(j,k) \in [0, \infty)$ *such that*

$$\lim_{N \to \infty} \boldsymbol{r}_N(j,k) = \boldsymbol{r}(j,k) .$$

Let $A \subset S$ *be the set of absorbing points of the Markovian dynamics induced by the rates* $\boldsymbol{r}(j,k)$. *Assume that for all* $j \in A$, $t > 0$,

$$\limsup_{N \to \infty} \max_{\xi \in \mathcal{E}_j^N} \mathbb{Q}_\xi^N \Big[\int_0^t \chi_{\Delta_N}(\xi_N(s)) \, ds \Big] = 0 .$$

Assume that for all $k \in S \setminus A$,

$$\lim_{N \to \infty} \frac{\pi_N(\Delta_N)}{\pi_N(\mathcal{E}_N^k)} = 0 .$$

Then, conditions (10.1), (10.2) *hold.*

This result, which guarantees tightness, together with Theorems 4.2, 5.1 and Remark 5.3, which provide uniqueness, yield the convergence of the sequence $\boldsymbol{X}_N^T$.

Theorem 10.4. *Fix* $k \in S$, *a sequence* $\eta_N \in \mathcal{E}_N^k$, *and denote by* $\mathbf{P}^N$ *the probability measure on* $D([0, \infty), S)$ *induced by the process* $\boldsymbol{X}_N^T(t)$ *and the measure* $\mathbb{P}_{\eta_N}^N$. *Assume the hypotheses of Theorem* 10.3. *Then, the sequence* $\mathbf{P}^N$ *converges to the solution of the* (L, δ_k) *martingale problem, where* L *is the generator of the* S-*valued Markov chain whose jump rates are* $\boldsymbol{r}(j,k)$.

11. The last passage

We prove in this section that the last passage process, introduced in Definition 2.1, converges if conditions (T1), (T2) hold. In order to prove this statement, we first define a metric in the path space $D([0, \infty), S \cup \{\mathfrak{d}\})$ which induces the Skorohod topology. Assume that $0 \notin S$ and identify the point $\mathfrak{d}$ with $0 \in \mathbb{Z}$ so that $S \cup \{\mathfrak{d}\}$ is a metric space with the metric induced by $\mathbb{Z}$.

For each integer $m \geq 1$, let Λ_m denote the class of strictly increasing, continuous mappings of $[0, m]$ onto itself. If $\lambda \in \Lambda_m$, then $\lambda_0 = 0$ and $\lambda_m = m$. In

addition, consider the function

$$
g_m(t) \;=\; \left\{ \begin{array}{ll} 1 & \text{if } \; t \le m-1 \,, \\ m-t & \text{if } \; m-1 \le t \le m \,, \\ 0 & \text{if } \; t \ge m \,. \end{array} \right.
$$

For any integer $m \ge 1$ and $\omega, \omega' \in D([0,\infty), S \cup \{\mathfrak{d}\})$, define $d_m(\omega, \omega')$ to be the infimum of those positive ϵ for which there exists $\lambda \in \Lambda_m$ satisfying

$$
\sup_{t \in [0,m]} |\lambda_t - t| \;<\; \epsilon \quad \text{and} \quad \sup_{t \in [0,m]} |\, g_m(\lambda_t)\, \omega(\lambda_t) \,-\, g_m(t)\, \omega'(t)\,| \;<\; \epsilon \,.
$$

Define the metric d in $D([0,\infty), S \cup \{\mathfrak{d}\})$ by

$$
d(\omega, \omega') \;=\; \sum_{m=1}^{\infty} \frac{1}{2^m} \left\{ 1 \wedge d_m(\omega, \omega') \right\} \,.
$$

This metric induces the Skorohod topology in the path space $D([0,\infty), S \cup \{\mathfrak{d}\})$ [26]. Next result is Proposition 4.4 in [15].

Recall from (2.3) the definition of $X_N^V(t)$ and let $\boldsymbol{X}_N^V(t) = X_N^V(t\theta_N)$. Recall from assumption (T1) the definition of $\boldsymbol{X}_N^T(t)$

Theorem 11.1. *Suppose that $(\eta_N(t) : t \ge 0)$, $N \ge 1$, satisfies condition (T2). Then, for any sequence $(\eta_N : N \ge 1)$, $\eta_N \in \mathcal{E}_N$,*

$$
\lim_{N \to \infty} \mathbb{E}_{\eta_N} \big[\, d(\boldsymbol{X}_N^V, \boldsymbol{X}_N^T)\,\big] \;=\; 0 \,.
$$

It follows from this result that the last-passage process $\boldsymbol{X}_N^V(t)$ converges whenever the trace process $\boldsymbol{X}_N^T(t)$ converges and (T2) is in force.

12. The finite-dimensional distributions

Recall the definition of the process $X_N(t)$ defined in (2.2), and the one of the reduced model $\boldsymbol{X}(t)$ introduced in Definition 2.1. Next result is Proposition 1.1 of [89].

Theorem 12.1. *Assume that conditions (T1) and (T2) of Definition 2.2 are in force, and that*

$$
\lim_{\delta \to 0} \limsup_{N \to 0} \sup_{\eta \in \mathcal{E}_N} \sup_{\delta \le s \le 2\delta} \mathbb{P}_\eta^N \big[\, \eta(s\theta_N) \in \Delta_N \,\big] \;=\; 0 \,. \tag{12.1}
$$

Then, the finite-dimensional distributions of $\boldsymbol{X}_N(t) = X_N(t\theta_N)$ converge to the finite-dimensional distributions of $\boldsymbol{X}(t)$.

With further mixing conditions one can prove that the state of the process at time $t\theta_N$ is a time-dependent convex combinations of states supported in the valleys.

Denote by $p_t(j, k)$ the transition probabilities of the reduced model $\boldsymbol{X}(t)$, by π_N^k the measure π_N conditioned to $\mathcal{E}_N^k$, and by $\|\mu - \nu\|_{\mathrm{TV}}$ the total variation distance between two probability measures μ and ν defined on E_N. Let $(S_N(t) : t \geq 0)$ be the semigroup associated to the Markov chain $\eta_N(t)$. Then, under mixing conditions specified in [89], for every $j \in S$ and sequence $\eta_N \in \mathcal{E}_j^N$,

$$\lim_{N \to \infty} \big\| \delta_{\eta_N}\, S_N(t\theta_N)\, -\, \sum_{k \in S} p_t(j, k)\, \pi_N^k \big\|_{\mathrm{TV}}\, =\, 0\, ,$$

where δ_η, $\eta \in E_N$, stands for the Dirac measure concentrated on the configuration η.

13. Examples

We present in this section some dynamics whose metastable behavior has been derived with the arguments presented in the article.

13.1. Random walks in a potential field

We describe the reversible version of the dynamics. The non-reversible one is obtained by replacing 2-cycles, in the terminology of Subsection A.2, by k-cycles.

Let Ξ be an open and bounded subset of $\mathbb{R}^d$, and denote by $\partial\Xi$ its boundary, which is assumed to be a smooth manifold. Fix a twice continuously differentiable function $F : \Xi \cup \partial\Xi \to \mathbb{R}$. We assume that the second partial derivatives of F are Lipschitz continuous; that all the eigenvalues of the Hessian of F at the critical points which are local minima are *strictly* positive; that the Hessian of F at the critical points which are not local minima or local maxima has one strictly negative eigenvalue, all the other ones being strictly positive. In dimension 1 this assumption requires the second derivative of F at the local minima to be strictly negative. Finally, we assume that for every $\boldsymbol{x} \in \partial\Xi$, $(\nabla F)(\boldsymbol{x}) \cdot \boldsymbol{n}(\boldsymbol{x}) > 0$, where $\boldsymbol{n}(\boldsymbol{x})$ represents the exterior normal to the boundary of Ξ, and $\boldsymbol{x} \cdot \boldsymbol{y}$ the scalar product of $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{R}^d$. This hypothesis guarantees that F has no local minima at the boundary of Ξ.

Denote by Ξ_N the discretization of Ξ: $\Xi_N = \Xi \cap (N^{-1}\mathbb{Z}^d)$, $N \geq 1$, where $N^{-1}\mathbb{Z}^d = \{\boldsymbol{k}/N : \boldsymbol{k} \in \mathbb{Z}^d\}$. The elements of Ξ_N are represented by the symbols $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_d)$, $\boldsymbol{y}$ and $\boldsymbol{z}$. Let μ_N be the probability measure on Ξ_N defined by

$$\mu_N(\boldsymbol{x})\, =\, \frac{1}{Z_N} e^{-NF(\boldsymbol{x})}\, , \quad \boldsymbol{x} \in \Xi_N\, ,$$

where Z_N is the partition function $Z_N = \sum_{\boldsymbol{x} \in \Xi_N} \exp\{-NF(\boldsymbol{x})\}$. Let $\{\eta_N(t) : t \geq 0\}$ be the continuous-time Markov chain on Ξ_N whose generator $\mathcal{L}_N$ is given by

$$(\mathcal{L}_N f)(\boldsymbol{x})\, =\, \sum_{\substack{\boldsymbol{y} \in \Xi_N \\ \|\boldsymbol{y} - \boldsymbol{x}\| = 1/N}} e^{-(1/2)N[F(\boldsymbol{y}) - F(\boldsymbol{x})]}\, [f(\boldsymbol{y}) - f(\boldsymbol{x})]\, , \qquad (13.1)$$

where $\| \cdot \|$ represents the Euclidean norm of $\mathbb{R}^d$. The rates were chosen for the measure μ_N to be reversible for the dynamics.

We restrict our atention here to the evolution among the shallowest valleys. One can infer from this discussion the general case which can be found in [91]. Denote by $\mathfrak{M}$ the set of local minima and by $\mathfrak{S}$ the set of saddle points of F in Ξ. Let $\mathfrak{S}_1$ be the set of the lowest saddle points:

$$\mathfrak{S}_1 \;=\; \left\{ \boldsymbol{z} \in \mathfrak{S} : F(\boldsymbol{z}) = \min\{F(\boldsymbol{y}) : \boldsymbol{y} \in \mathfrak{S}\} \right\} .$$

We represent by $\boldsymbol{z}^1, \ldots, \boldsymbol{z}^n$ the elements of $\mathfrak{S}_1$, $\mathfrak{S}_1 = \{\boldsymbol{z}^1, \ldots, \boldsymbol{z}^n\}$. Denote by H the height of the saddle points in $\mathfrak{S}_1$:

$$H \;=\; F(\boldsymbol{z}^1) .$$

Let $\widehat{\Omega}$ be the level set of Ξ defined by

$$\widehat{\Omega} \;=\; \left\{ \boldsymbol{x} \in \Xi : F(\boldsymbol{x}) \leq H \right\} .$$

The set $\widehat{\Omega}$ can be written as a disjoint union of connected components: $\widehat{\Omega} = \cup_{1 \leq j \leq \varkappa} \widehat{\Omega}_j$, where $\widehat{\Omega}_j \cap \widehat{\Omega}_k = \varnothing$, $j \neq k$, and where each set $\widehat{\Omega}_j$ is connected. Some connected component may not contain any saddle point in $\mathfrak{S}_1$, and some may contain more than one saddle point. Denote by Ω_j, $1 \leq j \leq m$, the connected components $\widehat{\Omega}_{j'}$ which contain a point in $\mathfrak{S}_1$.

Each component Ω_j is a union of valleys, $\Omega_j = W_{j,1} \cup \cdots \cup W_{j,m_j}$. The sets $W_{j,a}$ are defined as follows. Let $\mathring{\Omega}_j$ be the interior of Ω_j. Each set $W_{j,a}$ is the closure of a connected component of $\mathring{\Omega}_j$. The intersection of two valleys is a subset of the set of saddle points: $W_{j,a} \cap W_{j,b} \subset \mathfrak{S}_1$. Figure 6 illustrates the valleys of two connected components.
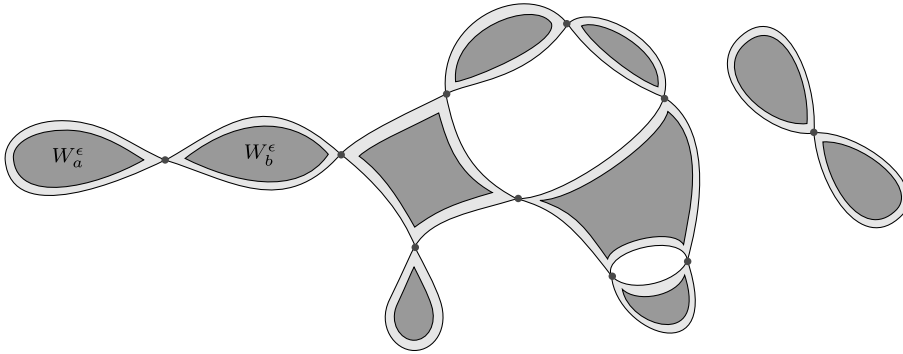

FIG 6. *Some valleys which form two connected components Ω_1 and Ω_2. The blue dots represent the saddle points and the gray regions the points $\boldsymbol{x}$ in the valleys such that $F(\boldsymbol{x}) < H - \epsilon$.*

Fix $1 \leq j \leq m$ and a connected component $\Omega = \Omega_j$. Let $S = \{1, \ldots, \ell\}$ denote the set of the indices of the valleys forming the connected component Ω: $\Omega = W_1 \cup \cdots \cup W_\ell$. Recall that $F(\boldsymbol{z}) = H$, $\boldsymbol{z} \in \mathfrak{S}_1$. For $\epsilon > 0$, $1 \leq a \leq \ell$, let

$W_a^\epsilon = \{\boldsymbol{x} \in W_a : F(\boldsymbol{x}) < H - \epsilon\}$, and let

$$\mathcal{E}_N^a \;=\; W_a^\epsilon \cap \Xi_N \;, \;\; 1 \le a \le \ell \;.$$

Each valley W_a contains exactly one local minimum of F, denoted by $\boldsymbol{m}_a$. Let $h_a = F(\boldsymbol{m}_a)$.

Let $\hat{\theta}_a = H - h_a > 0$, $a \in S$, be the depth of the valley W_a. The depths $\hat{\theta}_a$ provide the time-scale at which a metastable behavior is observed. Let $\theta_1 < \theta_2 < \cdots < \theta_p$, $p \le \ell$, be the increasing enumeration of the sequence $\hat{\theta}_a$, $1 \le a \le \ell$:

$$\{\hat{\theta}_1, \ldots, \hat{\theta}_\ell\} \;=\; \{\theta_1, \ldots, \theta_p\} \;.$$

The chain exhibits a metastable behavior on p different time scales in the set Ω. Let $T_q = \{a \in S : \hat{\theta}_a = \theta_q\}$, $1 \le q \le p$, so that $T_1, \ldots, T_p$ forms a partition of S, and let

$$S_q \;=\; T_q \cup \cdots \cup T_p \;, \;\; 1 \le q \le p \;.$$

Define the projection $\Psi_N^q : \Xi_N \to S_q \cup \{0\}$, $1 \le q \le p$, as

$$\Psi_N^q(\boldsymbol{x}) \;=\; \sum_{a \in S_q} a \, \chi_{\mathcal{E}_N^a}(\boldsymbol{x}) \;.$$

Note that $\Psi_N^q(\boldsymbol{x}) = 0$ for all points $\boldsymbol{x}$ which do not belong to $\cup_{a \in S_q} \mathcal{E}_N^a$. Denote by $X_N^q(t)$ the projection of the Markov chain $\eta_N(t)$ by Ψ_N^q:

$$X_N^q(t) \;=\; \Psi_N^q(X_N(t)) \;.$$

The theory presented in Sections 4–6 yields the existence, for each $1 \le q \le p$, of a time-scale β_N^q and a S_q-valued Markov chain $\boldsymbol{X}^q(t)$ with the following property. For each $a \in S_q$ and sequence of configurations $\boldsymbol{x}_N$ in $\mathcal{E}_N^a$, starting from $\boldsymbol{x}_N$, the finite-dimensional distributions of the projected process $\boldsymbol{X}_N^q(t) = X_N^q(t\beta_N^q)$ converge to the ones of $\boldsymbol{X}^q(t)$. The time-scales β_N^q can be explicitly computed and are related to the capacity between valleys.

We refer to [91, 92, 93, 90] for more details. This model is at the origin of the study of metastability from a dynamical point of view. The first results can be traced back at least to Hood [76], van't Hoff [75], Arrhenius [7], Eyring [58] and Kramers [83]. We refer to the recent books by Olivieri and Vares [111] and Bovier and den Hollander [31] and to the review by Berglund [23] for references and alternative derivations of these results.

13.2. Spin dynamics

Since the seminal paper by Cassandro, Galves, Olivieri and Vares [38], which introduced the pathwise approach to metastability, the metastable behavior of many spin dynamics have been derived in different ways. We do not review here the main results, but just illustrate the theory developed in the previous sections with one example. We again refer the reader to [111, 31] for a complete list of references on the subject.

Denote by $\Lambda_L = \{1, \ldots, L\}^2$ the two-dimensional discrete torus with L^2 elements, and let $\Omega_L = \{-1, 0, 1\}^{\Lambda_L}$. Elements of Ω_L are represented by the Greek letter σ. For $x \in \Lambda_L$, $\sigma(x) \in \{-1, 0, 1\}$ stands for the value at x of the configuration σ and is called the spin at x of σ.

The Blume–Capel model was introduced in [27, 36] to study the ${}^3\text{He}$ –${}^4\text{He}$ phase transition. One can think as a system of particles with spins. The value $\sigma(x) = 0$ corresponds to the absence of particles, while $\sigma(x) = \pm 1$ to the presence of a particle with spin equal to ± 1.

Fix an external field $h \in \mathbb{R}$, a magnetic field $\lambda \in \mathbb{R}$, and denote by $\mathbb{H} : \Omega_L \to \mathbb{R}$ the Hamiltonian given by

$$\mathbb{H}(\sigma) \;=\; \sum \left(\sigma(y) - \sigma(x)\right)^2 \;-\; h \sum_{x \in \Lambda_L} \sigma(x) \;-\; \lambda \sum_{x \in \Lambda_L} \sigma(x)^2 \,,$$

where the first sum is carried over all unordered pairs of nearest-neighbor sites of Λ_L.

Denote by μ_β the Gibbs measure associated to the Hamiltonian $\mathbb{H}$ at inverse temperature β. This is the probability measure on Ω_L given by

$$\mu_\beta(\sigma) = \frac{1}{Z_\beta} e^{-\beta \mathbb{H}(\sigma)}, \tag{13.2}$$

where Z_β is the partition function, the normalization constant which turns μ_β into a probability measure.

We refer to [47] for a description of the ground states, the configurations which minimize the Hamiltonian $\mathbb{H}$, according to the values of the parameters h and λ. In all cases, the ground states form a subset of the set $\{-\mathbf{1}, \mathbf{0}, +\mathbf{1}\}$, where $-\mathbf{1}, \mathbf{0}, +\mathbf{1}$ represent the configurations of Ω_L with all spins equal to $-1, 0, +1$, respectively.

The continuous-time Metropolis dynamics at inverse temperature β is the Markov chain on Ω_L, denoted by $\{\sigma_t : t \geq 0\}$, whose infinitesimal generator L_β acts on functions $f : \Omega_L \to \mathbb{R}$ as

$$(L_\beta f)(\sigma) = \sum_{x \in \Lambda_L} R_\beta(\sigma, \sigma^{x,+}) \left[f(\sigma^{x,+}) - f(\sigma)\right]$$
$$+ \sum_{x \in \Lambda_L} R_\beta(\sigma, \sigma^{x,-}) \left[f(\sigma^{x,-}) - f(\sigma)\right] .$$

In this formula, $\sigma^{x,\pm}$ represents the configuration obtained from σ by modifying the spin at x as follows,

$$\sigma^{x,\pm}(z) := \begin{cases} \sigma(x) \pm 1 \bmod 3 & \text{if } z = x \,, \\ \sigma(z) & \text{if } z \neq x \,, \end{cases}$$

where the sum is taken modulo 3, and the jump rates R_β are given by

$$R_\beta(\sigma, \sigma^{x,\pm}) \;=\; \exp\left\{ -\beta \left[\mathbb{H}(\sigma^{x,\pm}) - \mathbb{H}(\sigma)\right]_+ \right\}, \quad x \in \Lambda_L \,,$$

where a_+, $a \in \mathbb{R}$, stands for the positive part of a: $a_+ = \max\{a, 0\}$.

The Gibbs measure μ_β introduced in (13.2) satisfies the detailed balance conditions (A.3), and is therefore reversible for the dynamics.

Assume from now on that the chemical potential vanishes, $\lambda = 0$, and that the magnetic field h is small and positive, $0 < h < 2$. In this situation, the configurations $-\mathbf{1}$, $\mathbf{0}$ are local minima of the Hamiltonian, while the configuration $+\mathbf{1}$ is a global minimum. Moreover, $\mathbb{H}(\mathbf{0}) < \mathbb{H}(\mathbf{-1})$.

Assume that $2/h$ is not an integer and let $n_0 = \lfloor 2/h \rfloor$, where $\lfloor a \rfloor$ stands for the integer part of $a \in \mathbb{R}_+$. Denote by $\mathfrak{R}_c$ the set of configurations with $n_0(n_0 + 1) + 1$ 0-spins forming, in a background of -1-spins, a $n_0 \times (n_0 + 1)$ rectangle with an extra 0-spin attached to the longest side of this rectangle. This means that the extra 0-spin is surrounded by three -1-spins and one 0-spins which belongs to the longest side of the rectangle.

It is proved in [87, 88] that, as the temperature vanishes, starting from $-\mathbf{1}$ the process visits the set $\mathfrak{R}_c$ before hitting $\mathbf{0}$ or $+\mathbf{1}$:

$$\lim_{\beta \to \infty} \mathbb{P}_{-\mathbf{1}}[H_{\mathfrak{R}_c} < H_{\{\mathbf{0},+\mathbf{1}\}}] = 1 .$$

The set $\mathfrak{R}_c$ represents the energetic barrier which has to be surmounted to pass from $-\mathbf{1}$ to $\{\mathbf{0}, +\mathbf{1}\}$. Fix $\xi \in \mathfrak{R}_c$, let

$$\Delta = \mathbb{H}(\xi) - \mathbb{H}(-\mathbf{1}) = 4(n_0 + 1) - [n_0(n_0 + 1) + 1] ,$$

and let θ_β be given by

$$\theta_\beta = \frac{\mu_\beta(-\mathbf{1})}{\mathrm{cap}(-\mathbf{1}, \{\mathbf{0},+\mathbf{1}\})} = [1 + o_\beta(1)] \frac{3}{4(2n_0 + 1)} \frac{1}{|\Lambda_L|} e^{\Delta \beta} ,$$

where $o_\beta(1)$ is a remainder which vanishes as $\beta \to \infty$.

Fix $\mathfrak{d} \neq 0, \pm 1$, and denote by $\Psi : \Omega_L \to \{-1, 0, 1, \mathfrak{d}\}$ the projection defined by $\Psi(-\mathbf{1}) = -1$, $\Psi(\mathbf{0}) = 0$, $\Psi(+\mathbf{1}) = +1$, and $\Psi(\sigma) = \mathfrak{d}$, otherwise. The main results in [87, 88] state that, starting from $-\mathbf{1}$, the finite-dimensional distributions of the coarse-grained chain $\boldsymbol{X}_\beta(t) = \Psi(\sigma(\theta_\beta t))$ converge to the ones of the $\{-1, 0, 1\}$-valued, continuous-time Markov chain $\boldsymbol{X}(t)$ in which 1 is an absorbing state, and whose jump rates are given by

$$r(-1, 0) = r(0, 1) = 1 , \quad r(-1, 1) = r(0, -1) = 0 .$$

The metastable behavior of this model has been explored by Cirillo and Olivieri [47], Manzo and Olivieri [100], and more recently by Cirillo and Nardi [44], and Cirillo, Nardi and Spitoni [46]. The mean-field Potts model is another spin dynamics in which the spin may take more than two values. It has been examine recently in [92] and by Nardi and Zocca in [103].

13.3. Zero range processes

Denote by $\mathbb{N}$ the set of non-negative integers, $\mathbb{N} = \{0, 1, 2, ...\}$, by $\mathbb{T}_L$, $L \geq 1$, the discrete, one-dimensional torus with L points, and by η the elements of $\mathbb{N}^{\mathbb{T}_L}$

called configurations. The total number of particles at $x \in \mathbb{T}_L$ for a configuration $\eta \in \mathbb{N}^{\mathbb{T}_L}$ is represented by η_x. Let E_N, $N \geq 1$, be the set of configurations with N particles:

$$E_N := \Big\{ \eta \in \mathbb{N}^{\mathbb{T}_L} : \sum_{x \in \mathbb{T}_L} \eta_x = N \Big\} \, .$$

Fix $\alpha > 1$, and define $g : \mathbb{N} \to \mathbb{R}_+$ as

$$g(0) = 0 \, , \quad g(1) = 1 \quad \text{and} \quad g(n) = \frac{a(n)}{a(n-1)} \, , \quad n \geq 2 \, ,$$

where $a(0) = 1$, $a(n) = n^\alpha$, $n \geq 1$. In this way, $\prod_{i=1}^n g(i) = a(n)$, $n \geq 1$, and $\{g(n) : n \geq 2\}$ is a strictly decreasing sequence converging to 1 as $n \uparrow \infty$.

Fix $1/2 \leq p \leq 1$, and denote by $p(x)$ the transition probability given by $p(1) = p$, $p(-1) = 1-p$, $p(x) = 0$, otherwise. Let $\sigma^{x,y}\eta$ be the configuration obtained from η by moving a particle from x to y:

$$(\sigma^{x,y}\eta)_z = \begin{cases} \eta_x - 1 & \text{for } z = x \\ \eta_y + 1 & \text{for } z = y \\ \eta_z & \text{otherwise} \, . \end{cases}$$

The nearest-neighbor, zero-range process associated to the jump rates $\{g(k) : k \geq 0\}$ and the transition probability $p(x)$ is the continuous-time, E_N-valued Markov process $\{\eta_N(t) : t \geq 0\}$ whose generator L_N acts on functions $f : E_N \to \mathbb{R}$ as

$$(L_N f)(\eta) = \sum_{\substack{x,y \in \mathbb{T}_L \\ x \neq y}} g(\eta_x) \, p(y-x) \left\{ f(\sigma^{x,y}\eta) - f(\eta) \right\} \, .$$

Hence, if there are k particles at site x, at rate $pg(k)$, resp. $(1-p)g(k)$, one of them jumps to the right, resp. left. Since $g(k)$ decreases to 1 as $k \to \infty$, the more particles there are at some site x the slower they jump, but the rate remains bounded below by 1.

This Markov process is irreducible. The stationary probability measure, denoted by π_N, is given by

$$\pi_N(\eta) = \frac{N^\alpha}{Z_N} \prod_{x \in \mathbb{T}_L} \frac{1}{a(\eta_x)} \, ,$$

where Z_N is the normalizing constant.

Fix a sequence $\{\ell_N : N \geq 1\}$ such that $1 \ll \ell_N \ll N$, and let $\mathcal{E}_N^x$, $x \in \mathbb{T}_L$, be the set of configurations in which all but ℓ_N particles sit at x:

$$\mathcal{E}_N^x := \Big\{ \eta \in E_N : \eta_x \geq N - \ell_N \Big\} \, .$$

According to equation (3.2) in [15], for each $x \in \mathbb{T}_L$, $\pi_N(\mathcal{E}_N^x) \to 1/L$ as $N \uparrow \infty$.

Denote by $\eta^{\mathcal{E}_N}(t)$ the trace of the process $\eta_N(t)$ on $\mathcal{E}_N = \cup_x \mathcal{E}_N^x$, and let $\Psi_N : \mathcal{E}_N \mapsto S$ be given by

$$\Psi_N(\eta) = \sum_{x \in S} x \, \chi_{\mathcal{E}_N^x}(\eta) \, .$$

Under some further conditions on the sequence ℓ_N, it can be proven, following the method presented in Sections 4–6, that the time-rescaled coarse-grained process $\boldsymbol{X}_N(t) = X_N(tN^{1+\alpha}) = \Psi_N(\eta^{\mathcal{E}_N}(tN^{1+\alpha}))$ converges to a S-valued Markov chain $\boldsymbol{X}(t)$. The jump rates of the reduced model $\boldsymbol{X}(t)$ are proportional to the capacity of the random walk on the discrete torus with L points which jumps to the right with probability p and to the left with probability $1 - p$. Moreover, in the time scale $N^{1+\alpha}$ the time spent by the process $\eta_N(t)$ on $\Delta_N = E_N \setminus \mathcal{E}_N$ is negligible.

This model has been introduced by Evans [56] Godrèche examined the dynamics of the condensate in [68]. Its metastable behavior has been derived in [16, 85, 117]. The reduced model is a $\mathbb{T}_L$-valued Markov chain whose jump rates are proportional to the capacities of the underlying random walk associated to $p(\cdot)$.

The nucleation phase of this model has been described in [14]. Armendáriz, Grosskinsky and Loulakis [4] considered the case in which the total number of sites increases with the number of particles, keeping a constant density. In this situation, the reduced model is a Lévy-process.

Grosskinsky, Redig and Vafayi [72], Cao, Chleboun and Grosskinsky [34] and Bianchi, Dommers and Giardinà [25] proved the metastable behavior of the inclusion process, another interacting particle system which exhibits condensation.

Static aspects of condensation for this zero-range process and other dynamics have been examined by Jeon, March and Pittel [80], Grosskinsky, Schütz and Spohn [73], Armendáriz and Loulakis [5, 6], Chleboun and Grosskinsky [41, 42, 43], Grosskinsky, Redig and Vafayi [71], Godrèche and Luck [69], Armendáriz, Grosskinsky and Loulakis [3], Fajfrová, Gobron and Saada [59].

In some dynamics the condensate is formed instantaneously as the size of the system grows, Waclaw and Evans [123], Chau1, Connaughton and Grosskinsky [39].

13.4. *Random walks among random traps*

Let $(G_N : N \geq 1)$, $G_N = (V_N, E_N)$, be a sequence of possibly random, finite, connected graphs defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where V_N represents the set of vertices and E_N the set of unoriented edges. Assume that the number of vertices, $|V_N|$, converges to $+\infty$ in $\mathbb{P}$-probability. To fix ideas, one can consider the d-dimensional discrete torus with N^d points.

Assume that on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we are given an i.i.d collection of random variables $\{W_j^N : j \geq 1\}$, $N \geq 1$, independent of the random graph G_N and whose common distribution belongs to the basin of attraction of an α-stable law, $0 < \alpha < 1$. Hence, for all $N \geq 1$ and $j \geq 1$,

$$\mathbb{P}[W_j^N > t] = \frac{L(t)}{t^\alpha}, \quad t > 0,$$

where L is a slowly varying function at infinity.

For each $N \geq 1$, re-enumerate in decreasing order the weights $W_1^N, \ldots, W_{|V_N|}^N$: $\hat{W}_j^N = W_{\sigma(j)}^N$, $1 \leq j \leq |V_N|$ for some permutation σ of the set $\{1, \ldots, |V_N|\}$ and $\hat{W}_j^N \geq \hat{W}_{j+1}^N$ for $1 \leq j < |V_N|$. Let $(x_1^N, \ldots, x_{|V_N|}^N)$ be a random enumeration of the vertices of G_N and define $W_{x_j^N}^N = \hat{W}_j^N$, $1 \leq j \leq |V_N|$, turning $G_N = (V_N, E_N, W^N)$ into a finite, connected, vertex-weighted graph.

Consider for each $N \geq 1$, a continuous-time random walk $\{\eta_N(t) : t \geq 0\}$ on V_N, which waits a mean W_x^N exponential time at site x, after which it jumps to one of its neighbors with uniform probability. The generator $\mathcal{L}_N$ of this walk is given by:

$$(\mathcal{L}_N f)(x) = \frac{1}{\deg(x)} \frac{1}{W_x^N} \sum_{y \sim x} [f(y) - f(x)]$$

for every $f : V_N \to \mathbb{R}$, where $y \sim x$ means that $\{x, y\}$ belongs to the set of edges E_N and where $\deg(x)$ stands for the degree of x: $\deg(x) = \#\{y \in V_N : y \sim x\}$.

Let $\Psi_N : V_N \to \{1, \ldots, |V_N|\}$ be given by $\Psi_N(x_j^N) = j$. It has been proved for a class of random graphs that there exists a time-scale θ_N for which time-rescaled process $\boldsymbol{X}_N(t) = \Psi_N(\eta_N(t\theta_N))$ converges to a K-process.

To describe the dynamics of the K-process, consider two sequences of positive real numbers $\mathbf{u} = (u_k : k \geq 1)$ and $\mathbf{Z} = (Z_k : k \geq 1)$ such that

$$\sum_{k \geq 1} Z_k\, u_k \; < \; \infty \, , \quad \sum_{k \geq 1} u_k \; = \; \infty \, .$$

Consider the set $\overline{\mathbb{N}}_* = \{1, 2, \ldots\} \cup \{\infty\}$ of non-negative integers with an extra point denoted by ∞. We endow this set with the metric induced by the isometry $\phi : \overline{\mathbb{N}}_* \to \mathbb{R}$, which sends $n \in \overline{\mathbb{N}}_*$ to $1/n$ and ∞ to 0. This makes the set $\overline{\mathbb{N}}_*$ into a compact metric space.

The K-process with parameter (Z_k, u_k) can be informally described as follows. Being at $k \in \mathbb{N}$, the process waits a mean Z_k exponential time, at the end of which it jumps to ∞. Immediately after jumping to ∞, the process returns to $\mathbb{N}$. The hitting time of any finite subset A of $\mathbb{N}$ is almost surely finite. Moreover, for each fixed $n \geq 1$, the probability that the process hits the set $\{1, \ldots, n\}$ at the point k is equal to $u_k / \sum_{1 \leq j \leq n} u_j$. In particular, the trace of the K-process on the set $\{1, \ldots, n\}$ is the Markov process which waits at k a mean Z_k exponential time at the end of which it jumps to j with probability $u_j / \sum_{1 \leq i \leq n} u_i$.

In contrast with the theory presented in the previous sections, here the reduced model takes value in a countably infinite space. Moreover, as Ψ_N is a bijection, the process $\boldsymbol{X}_N(t)$ is Markovian, and we do not need to remove a piece of the state space by considering the trace, and we prove the convergence of the projection to the reduce model.

The K-process has been introduced by Fontes and Mathieu [63] who also proved the convergence to the K-process of the trap model in the complete graph. Fontes and Lima [62] considered the case of the hypercube. These results have been extended to d-dimensional torus, $d \geq 2$, and to random graphs in

[78, 79]. More recently, Cortines, Gold and Louidor considered a continuous time random walk on the two-dimensional discrete torus, whose motion is governed by the discrete Gaussian free field [49].

13.5. A polymer in the depinned phase

Fix $N \geq 1$ and denote by E_N the set of all lattice paths starting at 0 and ending at 0 after $2N$ steps:

$$E_N \;=\; \{\eta \in \mathbb{Z}^{2N+1} : \eta_{-N} = \eta_N = 0 \,, \; \eta_{j+1} - \eta_j = \pm 1 \,, \; -N \leq j < N\} \,.$$

Fix $0 < \alpha < 1$ and denote by $\eta_N(t)$ the E_N-valued Markov chain whose generator L_N is given by

$$(L_N f)(\eta) \;=\; \sum_{j=-N+1}^{N-1} c_{j,+}(\eta) \left[f(\eta^{j,+}) - f(\eta) \right] \;+\; \sum_{j=-N+1}^{N-1} c_{j,-}(\eta) \left[f(\eta^{j,-}) - f(\eta) \right] .$$

In this formula $\eta^{j,\pm}$ represents the configuration which is equal to η at every site $k \neq j$ and which is equal to $\eta_j \pm 2$ at site j.

The jump rate $c_{j,+}(\eta)$ vanishes at configurations η which do not satisfy the condition $\eta_{j-1} = \eta_{j+1} = \eta_j + 1$, and it is given by

$$c_{j,+}(\eta) \;=\; \begin{cases} 1/2 & \text{if } \eta_{j-1} = \eta_{j+1} \neq \pm 1, \\ 1/(1+\alpha) & \text{if } \eta_{j-1} = \eta_{j+1} = 1, \\ \alpha/(1+\alpha) & \text{if } \eta_{j-1} = \eta_{j+1} = -1 \end{cases}$$

for configurations which fulfill the condition $\eta_{j-1} = \eta_{j+1} = \eta_j + 1$. Let $-\eta$ stand for the configuration η reflected around the horizontal axis, $(-\eta)_j = -\eta_j$, $-N \leq j \leq N$. The rates $c_{j,-}(\eta)$ are given by $c_{j,-}(\eta) = c_{j,+}(-\eta)$.

Denote by $\Sigma(\eta)$ the number of zeros in the path η, $\Sigma(\eta) = \sum_{-N \leq j \leq N} \mathbf{1}\{\eta_j = 0\}$. The probability measure π_N on E_N defined by $\pi_N(\eta) = (1/Z_{2N})\, \alpha^{\Sigma(\eta)}$, where Z_{2N} is a normalizing constant, is easily seen to be reversible for the dynamics generated by L_N.

Denote by $\mathfrak{g}_N$ the spectral gap of the chain. The exact asymptotic behavior of $\mathfrak{g}_N$ is not known, but, by [37, Theorem 3.5], $\mathfrak{g}_N \leq C(\alpha)(\log N)^8/N^{5/2}$ for some finite constant $C(\alpha)$.

Fix a sequence ℓ_N such that $1 \ll \ell_N \ll N$, and let

$$\mathcal{E}_N^1 \;=\; \{\eta \in E_N : \eta_j > 0 \text{ for all } -(N - \ell_N) < j < (N - \ell_N)\} \,,$$
$$\mathcal{E}_N^2 = \{\eta \in E_N : -\eta \in \mathcal{E}_N^1\} \,, \quad \Delta_N \;=\; E_N \setminus (\mathcal{E}_N^1 \cup \mathcal{E}_N^2) \,.$$

By equation (2.27) in [35], $\pi_N(\mathcal{E}_N^1) = \pi_N(\mathcal{E}_N^1) = (1/2) + O(\ell_N^{-1/2})$.

Denote by $\mathfrak{g}_N^{R,j}$ the spectral gap of the chain reflected at $\mathcal{E}_N^j$, $j = 1, 2$. By [35, Proposition 2.6], taking $\ell_N = (\log N)^{1/4}$, for every $\epsilon > 0$, there exists N_0 such

that for all $N \geq N_0$, $\mathfrak{g}_N^{R,j} \geq N^{-(2+\epsilon)}$. In particular, choosing ϵ small enough and $\ell_N = (\log N)^{1/4}$,

$$\mathfrak{g}_N \ll \mathfrak{g}_N^{R,1}$$

for all N large enough. This shows that the chain equilibrates inside each valley in a much shorter time-scale than the one in which it jumps between valleys.

Let ν_N be a sequence of probability measures concentrated on $\mathcal{E}_N^1$ and which fulfills conditions (9.2). Set $\theta_N = 1/\mathfrak{g}_N$. The method presented in Section 9 yields that the time-rescaled coarse-grained process $\boldsymbol{X}_N^T(t) = X_N^T(t\theta_N)$, introduced in condition (T1) of Definition 2.2, converges to the $\{1,2\}$-valued Markov chain which starts from 1 and jumps from m to $3 - m$ at rate $1/2$. Moreover, in the time scale θ_N, the time spent by the process $\eta_N(t)$ outside the set $\mathcal{E}_N$ is negligible. We refer the reader to [20] for the proofs.

The interest of this model is that the entropy plays an important role. In contrast with the models presented in the previous subsections, the metastable behavior is not determined by an energy landscape, but by a repulsion in a bottleneck region of the space. In particular, in the terminology introduced in Remark 5.5, this dynamics does not visit points and the method presented in Sections 4–6 does not apply.

Note that the metastable behavior has been derived without a precise knowledge of the time-scale at which it occurs. Of course, the jumps between valleys take place in the time-scale θ_N, the inverse of the spectral gap, but the exact asymptotic behavior of $\mathfrak{g}_N$ is not known, and not needed in the proof of the metastable behavior of the dynamics.

This model has been introduced in [37, 35]. The results described in this subsection are taken from [20].

13.6. *Coalescing random walks*

Fix $d \geq 2$. Denote $\{e_1, \ldots, e_d\}$ the canonical basis of $\mathbb{R}^d$, and by p the probability measure on $\mathbb{Z}^d$ given by

$$p(x) \;=\; \frac{1}{2d} \text{ if } x \in \{\pm e_1, \ldots, \pm e_d\}\,, \quad p(x) = 0 \text{ otherwise} \,.$$

Let $\mathbb{T}_N^d$ be the discrete d-dimensional torus with N^d points. Denote by E_N the family of nonempty subsets of $\mathbb{T}_N^d$. Consider coalescing random walks on $\mathbb{T}_N^d$. This is the E_N-valued, continuous-time Markov chain, represented by $(A_N(t) : t \geq 0)$, whose generator L_N is given by

$$(L_N f)(A) = \sum_{x \in A} \sum_{y \notin A} p(y-x)\{f(A_{x,y}) - f(A)\} + \sum_{x \in A} \sum_{y \in A} p(y-x)\{f(A_x) - f(A)\},$$

where $A_{x,y}$, resp. A_x, is the set obtained from A by replacing the point x by y, resp. removing the element x:

$$A_{x,y} \;=\; [A \setminus \{x\}] \cup \{y\}\,, \quad A_x \;=\; A \setminus \{x\} \,.$$

In contrast with the previous dynamics, in this example the reduced model takes value in a countably infinite state space. Let $S = \{1, 1/2, 1/3, \dots\} \cup \{0\}$, and let $C^1(S)$ be the set of functions $f : S \to \mathbb{R}$ of class C^1, that is $f \in C^1(S)$ is the restriction to S of a continuously differentiable function defined on $\mathbb{R}$. For each $f \in C^1(S)$ define $\mathcal{L}f : S \to \mathbb{R}$ as

$$(\mathcal{L}f)(y) := \begin{cases} \binom{n}{2}\left\{f\left(\frac{1}{n-1}\right) - f\left(\frac{1}{n}\right)\right\}, & \text{if } y = \frac{1}{n} \text{ and } n \geq 2 \,, \\ 0 \,, & \text{if } y = 1 \,, \\ (1/2)f'(0) \,, & \text{if } y = 0 \,. \end{cases}$$

Proposition 2.1 in [13] asserts that for each $x \in S$ there exists a unique solution to the $(\mathcal{L}, \delta_x)$-martingale problem.

Consider the partition of E_N given by

$$E_N = \bigcup_{n \in \mathbb{N}} \mathcal{E}_N^n \,, \quad \text{where} \quad \mathcal{E}_N^n := \{A \subset \mathbb{T}_N^d : |A| = n\} \,, \quad n \in \mathbb{N} \,.$$

In this formula, $|A|$ stands for the number of elements of A. Let $\Psi_N : E_N \to S$ be the corresponding projection:

$$\Psi_N(A) = 1/|A| \,, \quad A \in E_N \,.$$

To define the metastable time-scale, consider two independent random walks $(x_t^N)_{t \geq 0}$ and $(y_t^N)_{t \geq 0}$ on $\mathbb{T}_N^d$, both with jump probability given by $p(\cdot)$, starting at the uniform distribution. Let θ_N be the expected meeting time:

$$\theta_N := E\left[\min\{t \geq 0 : x_t^N = y_t^N\}\right] . \tag{13.3}$$

Since $x_t^N - y_t^N$ evolves as a random walk speeded-up by 2, θ_N represents the expectation of the hitting time of the origin for a simple symmetric random walk speeded-up by 2 which starts from the uniform measure. In a general graph, though, the time-scale should be given by (13.3) mutatis mutandis.

Consider a continuous-time, random walk $(x_t)_{t \geq 0}$ on $\mathbb{Z}^d$ with jump probabilities given by $p(\cdot)$ and which starts from the origin. Assume that $d \geq 3$, and denote by v_d the escape probability: $v_d = P_0[H_0^+ = \infty]$. It can be shown that

$$\lim_{N \to \infty} \frac{\theta_N}{N^d} = \frac{1}{2\,v_d} \quad \text{in dimension } d \geq 3 \,,$$

$$\lim_{N \to \infty} \frac{\theta_N}{N^2 \log N} = \frac{1}{\pi} \quad \text{in dimension } d = 2 \,.$$

The factor 2 in the denominator appears because the process has been speeded-up by 2. In particular, in $d = 2$, $1/\pi$ should be understood as $(1/2)(2/\pi)$. We refer to [13] for a proof of this result.

Consider the time-rescaled coarse-grained process

$$\mathbb{X}_N(t) = \Psi_N(A_N(\theta_N t)) \,, \quad t \geq 0 \,.$$

Note that in this example we do not take the trace of the process on some set, but we just project it on a smaller state space.

Applying the ideas presented in the previous sections, it is proved in [13] that, starting from the configuration in which each site is occupied by a particle, $\mathbb{X}_N(t)$ converges in the Skorohod topology to the Markov chain whose generator is given by $\mathcal{L}$ and which starts from 0.

This model has been first considered by Cox [50], who proved that the coalescence time [the time all particles coalesced into one] is asymptotically equal to a sum of independent exponential random variables. This result has been extended by Oliveira [107, 108] to the case of transitive graphs. Related questions have been examined by Aldous and Fill [2], Durrett [53], Cooper, Frieze and Radzik [48], Chen, Choi and Cox [40].

13.7. *Further examples*

We mention in this last subsection other models whose metastable behavior has been derived with the tools presented in the previous sections.

The metastable behavior of sequences of continuous-time Markov chains on a fixed finite state-space has been examined in [17, 95]. This problem has been addressed with large deviations techniques by Scopolla [116], Olivieri and Scopolla in [109, 110], Manzo, Nardi, Olivieri and Scoppola [99] and Cirillo, Nardi and Sohier [45].

Properties of hitting times of rare events have been considered in [22]. Fernandez, Manzo, Nardi, Scoppola and Sohier [60], and Fernandez, Manzo, Nardi and Scoppola [61] examined this question through the pathwise approach.

The evolution, in the zero-temperature limit, of a droplet in the Ising model under the conservative Kawasaki dynamics in a large two-dimensional square with periodic boundary conditions has been derived in [18, 70]. The reduced model in this example is a two-dimensional Brownian motion on the torus.

Misturini [102] considered the ABC model on a ring in a strongly asymmetric regime. He derived the metastable behavior of the dynamics among the segregated configurations in the zero-temperature limit. Here, the reduced model is a Brownian motion.

Appendix A: Markov chains

We briefly present in this section some results on Markov chains used in the article. Fix a finite set E. Consider a continuous-time, E-valued, Markov chain $(\eta(t) : t \geq 0)$. Assume that the chain $\eta(t)$ is irreducible and denote by π the unique stationary state.

Elements of E are represented by the letters η, ξ. Let $\mathbb{P}_\eta$, $\eta \in E$, be the probability measure on $D([0, \infty), E)$ induced by the Markov chain $\eta(t)$ starting from η. Recall from (2.1) the definition of the hitting time and the return time to a set.

Denote by $R(\eta, \xi)$, $\eta \neq \xi \in E$, the jump rates of the Markov chain $\eta(t)$, and let $\lambda(\eta) = \sum_{\xi \in E} R(\eta, \xi)$ be the holding rates. Denote by $p(\eta, \xi)$ the jump probabilities, so that $R(\eta, \xi) = \lambda(\eta) \, p(\eta, \xi)$. The stationary state of the embedded discrete-time Markov chain is given by $M(\eta) = \pi(\eta) \, \lambda(\eta)$.

Denote by L the generator of the Markov chain $\eta(t)$,

$$(Lf)(\eta) \; = \; \sum_{\xi \in E} R(\eta, \xi) \, [\, f(\xi) - f(\eta) \,] \; .$$

Let $L^2(\pi)$ be the set of square-summable functions $f : E \to \mathbb{R}$ endowed with the scalar product $\langle \, \cdot \, , \, \cdot \, \rangle_\pi$ given by

$$\langle \, f \, , \, g \, \rangle_\pi \; := \; \sum_{\eta \in E_N} f(\eta) \, g(\eta) \, \pi(\eta) \; , \quad \| \, f \, \|^2 \; = \; \langle \, f \, , \, f \, \rangle_\pi \; .$$

Denote by L^* the adjoint of the operator L in $L^2(\pi)$: For all functions f, g : $E \to \mathbb{R}$,

$$\langle \, L^* f \, , \, g \, \rangle_\pi \; = \; \langle \, f \, , \, L \, g \, \rangle_\pi \; . \tag{A.1}$$

An elementary computation yields that

$$(L^* f)(\eta) \; = \; \sum_{\xi \in E} R^*(\eta, \xi) \, [\, f(\xi) - f(\eta) \,] \; ,$$

where the jump rates $R^*(\eta, \xi)$ satisfy

$$\pi(\eta) \, R^*(\eta, \xi) \; = \; \pi(\xi) \, R(\xi, \eta) \; , \quad \eta \neq \xi \in E \; . \tag{A.2}$$

The chain is said to be *reversible* if the generator L is self-adjoint: $L^* = L$. It is reversible if and only if the jump rates satisfy the *detailed balance conditions*:

$$\pi(\eta) \, R(\eta, \xi) \; = \; \pi(\xi) \, R(\xi, \eta) \; , \quad \eta \neq \xi \in E \; . \tag{A.3}$$

The operator L^* corresponds to the generator of a Markov chain, represented by $\eta^*(t)$, and called the adjoint or time-reversed process. The holding rates $\lambda^*(\eta) = \sum_{\xi \in E} R^*(\eta, \xi)$ of this chain coincide with the original ones, $\lambda^*(\eta) = \lambda(\eta)$, and the jump probabilities $p^*(\eta \, \xi)$ satisfy the balance conditions

$$M(\eta) \, p^*(\eta, \xi) \; = \; M(\xi) \, p(\xi, \eta) \; , \quad \eta \neq \xi \in E \; . \tag{A.4}$$

Let L^s be the symmetric part of the generator L:

$$L^s \; = \; \frac{1}{2} \, \{ \, L \, + \, L^* \, \} \; . \tag{A.5}$$

The operator L^s is self-adjoint in $L^2(\pi)$ and it corresponds to the generator of the Markov chain whose jump rates, denoted by $R^s(\eta, \xi)$, are given by $R^s(\eta, \xi) = (1/2)\{R(\eta, \xi) + R^*(\eta, \xi)\}$. A simple computation shows that these rates satisfy the detailed balance conditions (A.3).

Denote by $D(f)$ the Dirichlet form of a function $f : E \to \mathbb{R}$:

$$D(f) \; := \; \langle\, (-L)\, f\,,\, f\, \rangle_\pi \; = \; \langle\, (-L^s)\, f\,,\, f\, \rangle_\pi \,. \tag{A.6}$$

We leave to the reader the assignment of checking the last equality. An elementary computation shows that

$$D(f) \; = \; \frac{1}{2} \sum_{\eta \in E} \sum_{\xi \in E} \pi(\eta)\, R(\eta, \xi)\, [\, f(\xi) - f(\eta)\,]^2 \,. \tag{A.7}$$

This formula holds even in the non-reversible case. In the sum, each unordered pair $\{\eta, \xi\} \subset E$, $\xi \neq \eta$, appears twice.

Denote by $(S(t) : t \geq 0)$, the semigroup associated to the generator L, so that $(d/dt)S(t) = L\, S(t) = S(t)\, L$. Fix a probability measure ν on E and let f_t be the Radon-Nikodym derivative of $\nu S(t)$ with respect to π. We claim that

$$\frac{d}{dt} f_t \; = \; L^* f_t \,. \tag{A.8}$$

Indeed, fix a function $g : E \to \mathbb{R}$ and consider the mean $\mathbb{E}_\nu[g(\eta(t))]$, where $\mathbb{E}_\nu$ represents the expectation with respect to the measure $\mathbb{P}_\nu = \sum_{\eta \in E} \nu(\eta)\, \mathbb{P}_\eta$. This expectation can be written as

$$\sum_{\eta \in E} \nu(\eta)\, [S(t)\, g](\eta) \; = \; \sum_{\eta \in E} [\nu\, S(t)](\eta)\, g(\eta) \; = \; \sum_{\eta \in E} \pi(\eta)\, f_t(\eta)\, g(\eta) \; = \; \langle\, f_t\,,\, g\, \rangle_\pi \,.$$

As $(d/dt)S(t)g = S(t)\, L\, g$, taking derivative on both sides of this identity we get that

$$\sum_{\eta \in E} \nu(\eta)\, [S(t)\, L\, g](\eta) \; = \; \langle\, \frac{d}{dt}\, f_t\,,\, g\, \rangle_\pi \,.$$

The left-hand side can be written as $\langle\, f_t\,,\, L\, g\, \rangle_\pi \; = \; \langle\, L^* f_t\,,\, g\, \rangle_\pi$. Hence, for all functions g, $\langle\, (d/dt)f_t\,,\, g\, \rangle_\pi \; = \; \langle\, L^* f_t\,,\, g\, \rangle_\pi$, which proves claim (A.8).

By (A.8) and (A.1),

$$\frac{d}{dt}\, \langle\, f_t\,,\, f_t\, \rangle_\pi \; = \; 2\, \langle\, L\, f_t\,,\, f_t\, \rangle_\pi \; = \; -\, 2\, D(f_t) \; \leq \; 0 \,.$$

The inequality follows from the positiveness of the Dirichlet form derived in (A.7). Integrating in time yields that

$$\|\, f_t\, \|^2 \; + \; 2 \int_0^t D(f_s)\, ds \; \leq \; \|\, f_0\, \|^2 \,.$$

In particular, for all $0 \leq s \leq t$,

$$\langle\, f_t\,,\, f_t\, \rangle_\pi \; \leq \; \langle\, f_s\,,\, f_s\, \rangle_\pi \,. \tag{A.9}$$

The *spectral gap* of the generator, denoted by $\mathfrak{g}$, is the value of the smallest positive eigenvalue of the symmetric part of the generator:

$$\mathfrak{g} \; = \; \inf_f \frac{\langle\, (-L)\, f\,,\, f\, \rangle_\pi}{\langle\, f\,,\, f\, \rangle_\pi} \,,$$

where the infimum is carried over all functions $f : E \to \mathbb{R}$ which are orthogonal to the constants, i.e., which have zero-mean with respect to π: $E_\pi[f] = \langle\, f\, ,\, 1\,\rangle_\pi = 0$.

A.1. Reflected chain

Fix a non-empty, proper subset F of E. Denote by $(\eta^{R,F}(t) : t \geq 0)$, the Markov chain $\eta(t)$ reflected at F. This is the F-valued process obtained from $\eta(t)$ by forbidding all jumps between F and $E \setminus F$. The generator $L_{R,F}$ of this Markov process is given by

$$(L_{R,F}f)\,(\eta) \;=\; \sum_{\xi \in F} R(\eta, \xi) \left\{ f(\xi) - f(\eta) \right\} , \quad \eta \in F\ .$$

Assume that the reflected process $\eta^{R,F}(t)$ is irreducible. It is easy to show that the conditioned probability measure π_F defined by

$$\pi_F(\eta) \;=\; \frac{\pi(\eta)}{\pi(F)}\ , \quad \eta \in F\ , \tag{A.10}$$

satisfies the detailed balance conditions (A.3) for the reflected process if the chain is reversible.

In general, π_F may not be invariant. Consider, for example, an asymmetric random walk on the circle. The uniform measure is invariant, but its restriction to an interval I is not invariant for the process reflected at I. For cycle generators, however, it is possible to reflect the chain preserving the stationary state.

A.2. Cycle generators

The results of this subsection are taken from Section 4 of [95]. We refer to [93] for an application.

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

Cycle generator: A generator L is said to be a cycle generator associated to the cycle $\mathfrak{c} = (\eta_0, \eta_1, \ldots, \eta_{n-1}, \eta_n = \eta_0)$ if there exists reals $r_i > 0, 0 \leq i < n$, such that

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

We denote this cycle generator by $\mathcal{L}_{\mathfrak{c},\boldsymbol{r}}$, where $\boldsymbol{r} = (r_0, \ldots, r_{n-1})$. Most of the time we omit the dependence on $\boldsymbol{r}$ and write $\mathcal{L}_{\mathfrak{c},\boldsymbol{r}}$ simply as $\mathcal{L}_{\mathfrak{c}}$. Note that

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

and that the chain is irreducible only if $\{\eta_0, \eta_1, \ldots, \eta_{n-1}\} = E$.

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

$$\pi(\eta_i)\, R(\eta_i, \eta_{i+1}) \ \text{is constant .} \tag{A.11}$$

Sector condition: Next lemma asserts that every cycle generator satisfies a sector condition. The proof of this result can be found in [82, Lemma 5.5.8].

Lemma A.1. *Let $\mathcal{L}_\mathfrak{c}$ be a cycle generator associated to a cycle $\mathfrak{c}$ of length n. Then, $\mathcal{L}_\mathfrak{c}$ satisfies a sector condition with constant $2n$: For all $f,\, g : E \to \mathbb{R}$,*

$$\langle\, \mathcal{L}_\mathfrak{c}\, f\, ,\, g\, \rangle_\pi^2 \ \leq \ 2n\, \langle\, (-\,\mathcal{L}_\mathfrak{c}\, f)\, ,\, f\, \rangle_\pi \, \langle\, (-\,\mathcal{L}_\mathfrak{c}\, g)\, ,\, g\, \rangle_\pi\, .$$

Cycle decomposition: Every generator L, stationary with respect to a probability measure π, can be decomposed as the sum of cycle generators which are stationary with respect to π.

Lemma A.2. *Let L be a generator of an E-valued, irreducible Markov chain. Denote by π the unique invariant probability measure. Then, there exists cycles $\mathfrak{c}_1, \ldots, \mathfrak{c}_p$ such that*

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

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

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 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) coincides with the cycle (ξ, η, ξ).

Fix a cycle $\mathfrak{c} = (\eta, \xi, \eta) \in \mathbb{C}_2$. Let $\bar{c}(\eta, \xi) = \min\{\pi(\eta)R(\eta, \xi), \pi(\xi)R(\xi, \eta)\}$ be the minimal conductance of the edge (η, ξ), and let $R_\mathfrak{c}(\eta, \xi)$ be the jump rates given by $R_\mathfrak{c}(\eta, \xi) = \bar{c}(\eta, \xi)/\pi(\eta)$, $R_\mathfrak{c}(\xi, \eta) = \bar{c}(\eta, \xi)/\pi(\xi)$. 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 $\pi(\eta)R_\mathfrak{c}(\eta, \xi) = \bar{c}(\eta, \xi) = \pi(\xi)R_\mathfrak{c}(\xi, \eta)$, by (A.11), π is a stationary state for $\mathcal{L}_\mathfrak{c}$ (actually, reversible). Let $\mathcal{L}_1 = \mathcal{L} - \mathcal{L}_\mathfrak{c}$ so that

$$L \ = \ L_1 \ + \ \mathcal{L}_\mathfrak{c}\, .$$

As $R_\mathfrak{c}(\zeta, \zeta') \leq R(\zeta, \zeta')$, L_1 is the generator of a Markov chain. Since both L and $\mathcal{L}_\mathfrak{c}$ are stationary for π, so is 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 L_1 is equal to the number of arrows for the generator L minus 1 or 2. This procedure has therefore strictly decreased the number of arrows of L.

We may repeat the previous algorithm to L_1 to remove from 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 $\prod_{0 \le i < 3} R(\eta_i, \eta_{i+1}) > 0$. 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, $2 \le k < |E|$, we have obtained a decomposition of L as

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

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

Corollary A.3. *The generator L satisfies a sector condition with constant bounded by $2|E|$: For all f, $g : E \to \mathbb{R}$,*

$$\langle\, L f\,,\, g\,\rangle_\pi^2 \ \le \ 2|E|\,\langle\,(-\,L f)\,,\, f\,\rangle_\pi\,\langle\,(-\,L g)\,,\, g\,\rangle_\pi\,.$$

Proof. Fix f and $g : E \to \mathbb{R}$. By Lemma A.2,

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

where $\mathcal{L}_{\mathfrak{c}_j}$ is a cycle generator, stationary with respect to π, associated to the cycle $\mathfrak{c}_j$. By Lemma A.1 and by Schwarz inequality, since all cycles have length at most $|E|$, the previous sum is bounded by

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

as claimed $\qquad\qquad \square$

Remark A.4. *A generator L is reversible with respect to π if and only if it has a decomposition in 2-cycles. Given a measure π on a finite state space, by introducing k-cycles satisfying (A.11) it is possible to define non-reversible dynamics which are stationary with respect to π. The previous lemma asserts that this is the only way to define such dynamics.*

Remark A.5. *The decomposition in cycles is not unique. There may exist cycles and vectors $\mathfrak{c}_1, \ldots, \mathfrak{c}_p$, $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_p$ and $\hat{\mathfrak{c}}_1, \ldots, \hat{\mathfrak{c}}_q$, $\hat{\boldsymbol{r}}_1, \ldots, \hat{\boldsymbol{r}}_q$ such that $\{\mathfrak{c}_1, \ldots, \mathfrak{c}_p\} \ne \{\hat{\mathfrak{c}}_1, \ldots, \hat{\mathfrak{c}}_q\}$,*

$$L = \sum_{j=1}^{p} \mathcal{L}_{\mathfrak{c}_j, \boldsymbol{r}_j} \ = \ \sum_{k=1}^{q} \mathcal{L}_{\hat{\mathfrak{c}}_k, \hat{\boldsymbol{r}}_k}\,,$$

and π is a stationary state for all cycle generators. We leave the reader to find an example. However, in view of Lemma A.1, it is natural to look for one which minimizes the length of the longest cycle.

Remark A.6. *In a finite set, the decomposition of a generator into cycle generators is very simple. The problem for countably-infinite sets is much more delicate. We refer to [65] for a discussion.*

Let F be a proper subset of E and consider the chain reflected at F. The last result of this subsection provides sufficient conditions for the measure π conditioned to F to be a stationary state for the reflected process in the non-reversible case.

Lemma A.7. *Assume that the generator L can be written as a sum of cycle generators:*

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

where $\mathfrak{c}_1, \ldots, \mathfrak{c}_p$ are cycles and π is a stationary state for each $\mathcal{L}_{\mathfrak{c}_j}$. Then, the measure π conditioned to F is stationary for the reflected chain at F if there exists a subset A of $\{1, \ldots, p\}$ such that

$$L_{R,F} \;=\; \sum_{j \in A} \mathcal{L}_{\mathfrak{c}_j} \;.$$

Proof. Since π is a stationary state for each $\mathcal{L}_{\mathfrak{c}_j}$, it is also a stationary state for $L_{R,F} = \sum_{j \in A} \mathcal{L}_{\mathfrak{c}_j}$. As the reflected process does not leave the set F, the measure π is stationary if and only if its restriction to F is stationary. $\qquad\square$

A.3. Enlarged chains

Let $E^\star$ be a copy of E. The elements of $E^\star$ are represented by the letters η, ξ. Denote by $P_\star : E \cup E^\star \to E \cup E^\star$ the application which maps a configuration in E, $E^\star$, to its copy in $E^\star$, E, respectively.

Following [24], for $\gamma > 0$ denote by $\eta^\gamma(t)$ the Markov process on $E \cup E^\star$ whose jump rates $R^\gamma(\eta, \xi)$ are given by

$$R^\gamma(\eta, \xi) \;=\; \begin{cases} R(\eta, \xi) & \text{if } \eta \text{ and } \xi \in E, \\ 1/\gamma & \text{if } \xi = P_\star \eta, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, being at some state ξ in $E^\star$, the process may only jump to $P_\star \xi$ and this happens at rate $1/\gamma$. In contrast, being at some state ξ in E, the process $\eta^\gamma(t)$ jumps with rate $R(\xi, \xi')$ to the state $\xi' \in E$, and jumps with rate $1/\gamma$ to $P_\star \xi$. We call the process $\eta^\gamma(t)$ the *γ-enlargement* of the process $\eta(t)$.

Let $\pi_\star$ be the probability measure on $E \cup E^\star$ defined by

$$\pi_\star(\eta) \;=\; \pi_\star(P_\star \eta) \;=\; (1/2)\,\pi(\eta) \;, \quad \eta \in E \;.$$

The probability measure $\pi_\star$ is invariant for the enlarged process $\eta^\gamma(t)$ and it is reversible whenever π is reversible.

Let F be a subset of E. Think of F as a valley. If γ is much larger than the mixing time, the distribution of $\eta(H_{F_\star})$, where $F_\star = \{P_\star \eta : \eta \in F\}$, is very close the stationary state conditioned to F.

A.4. *Collapsed chains*

The collapsed chain consists in collapsing a subset of the state-space to a point and in the defining a dynamics which keeps the properties of the original evolution as much as possible. This is a well-known technique, see for instance [32, 1].

Fix a subset A of E, and let $E_A := [E \setminus A] \cup \{\mathfrak{d}\}$, where $\mathfrak{d}$ stands for an extra configuration added to E and meant to represent the collapsed set A. Denote by $(\eta^{C,A}(t) : t \geq 0)$ the chain obtained from $\eta(t)$ by collapsing the set A to the singleton $\{\mathfrak{d}\}$. This is the continuous-time Markov chain on E_A with jump rates $R^{C,A}(\eta, \xi)$, $\eta, \xi \in E_A$, given by

$$R^{C,A}(\eta, \xi) \;=\; R(\eta, \xi)\,, \quad R^{C,A}(\eta, \mathfrak{d}) \;=\; \sum_{\zeta \in A} R(\eta, \zeta)\,, \quad \eta\,,\, \xi \,\in\, E \setminus A\,,$$

$$R^{C,A}(\mathfrak{d}, \eta) \;=\; \frac{1}{\pi(A)} \sum_{\xi \in A} \pi(\xi)\, R(\xi, \eta)\,, \quad \eta \in E \setminus A\,.$$

(A.12)

The collapsed chain $\{\eta^{C,A}(t) : t \geq 0\}$ inherits the irreducibility from the original chain. Denote by $\pi^{C,A}$ the probability measure on E_A given by

$$\pi^{C,A}(\mathfrak{d}) \;=\; \pi(A)\,, \quad \pi^{C,A}(\eta) \;=\; \pi(\eta)\,, \quad \eta \in E \setminus A\,. \qquad \text{(A.13)}$$

Since

$$\sum_{\xi \notin A, \zeta \in A} \pi(\xi)\, R(\xi, \zeta) \;=\; \sum_{\xi \notin A, \zeta \in A} \pi(\zeta)\, R(\zeta, \xi)\,,$$

one checks that $\pi^{C,A}$ is a stationary state, and therefore the unique invariant probability measure, for the collapsed chain $\eta^{C,A}(t)$.

The collapsed chain has to be understood as follows. Until the process hits the set A, it evolves as the original one. When it reaches this set, it immediately equilibrates and its position is replaced by the stationary distribution conditioned to A.

In particular, we may couple the collapsed process with the original one until the set A is reached, so that, for every $\eta \in E \setminus A$, and $B \subset E \setminus A$,

$$\mathbb{P}^{C,A}_\eta\big[\, H_\mathfrak{d} < H_B^+ \,\big] \;=\; \mathbb{P}_\eta\big[\, H_A < H_B^+ \,\big]\,, \qquad \text{(A.14)}$$

provided $\mathbb{P}^{C,A}_\eta$ represents the distribution of the collapsed chain $\eta^{C,A}(t)$ starting from η. It follows from this identity and the explicit formulae for the jump rates and the stationary state that for every $B \subset E \setminus A$,

$$\mathrm{cap}(A, B) \;=\; \mathrm{cap}^{C,A}(\mathfrak{d}, B)\,,$$

where $\mathrm{cap}^{C,A}(\mathfrak{d}, B)$ represents the capacity between $\mathfrak{d}$ and B for the collapsed chain.

This identity ceases to hold if we replace A by a set in $E \setminus A$ because (A.14) is incorrect if $\mathfrak{d}$, A are replaced by a set $D \subset E \setminus A$.

Denote by $L^{C,A}$ the generator of the chain $\eta^{C,A}(t)$. Fix two functions f, $g : E_A \to \mathbb{R}$. Let F, $G : E \to \mathbb{R}$ be defined by

$$F(\eta) \;=\; f(\eta)\,, \quad \eta \in E \setminus A\,, \quad F(\zeta) = f(\mathfrak{d})\,, \quad \zeta \in A\,,$$

with a similar definition for G. We claim that

$$\langle\, L^{C,A} f\,,\, g\,\rangle_{\pi^{C,A}} \;=\; \langle\, LF\,,\, G\,\rangle_\pi \,. \tag{A.15}$$

Conversely, if F, $G : E \to \mathbb{R}$ are two functions constant over A, (A.15) holds if we define f, $g : E_A \to \mathbb{R}$ by

$$f(\eta) \;=\; F(\eta)\,, \quad \eta \in E \setminus A\,, \quad f(\mathfrak{d}) \;=\; F(\zeta) \quad \text{for some} \quad \zeta \in A\,,$$

with an analogous equation for f, F replaced by g, G, respectively.

To prove (A.15), fix two functions f, $g : E_A \to \mathbb{R}$. By definition of $L^{C,A}$,

$$\langle\, L^{C,A} f\,,\, g\,\rangle_{\pi^{C,A}} \;=\; \sum_{\eta,\xi \in E_A} \pi^{C,A}(\eta)\, R^{C,A}(\eta,\xi)\, [\, f(\xi) - f(\eta)\,]\, g(\eta)\,.$$

In view of (A.12), (A.13), this expression is equal to

$$\sum_{\eta \in E \setminus A} \pi(\eta) \Big\{ \sum_{\xi \in E \setminus A} R(\eta,\xi)\, [\, f(\xi) - f(\eta)\,] \;+\; \sum_{\zeta \in A} R(\eta,\zeta)\, [\, f(\mathfrak{d}) - f(\eta)\,] \Big\} g(\eta)$$
$$+\; \sum_{\xi \in E \setminus A}\sum_{\zeta \in A} \pi(\zeta)\, R(\zeta,\xi)\, [\, f(\xi) - f(\mathfrak{d})\,]\, g(\mathfrak{d})\,.$$

Since $F(\eta) = f(\eta)$ for $\eta \in E \setminus A$, and $F(\xi) = f(\mathfrak{d})$ for $\xi \in A$, with similar identities with G, g replacing F, f, the last sum is equal to

$$\sum_{\eta \in E \setminus A} \pi(\eta) \Big\{ \sum_{\xi \in E \setminus A} R(\eta,\xi)\, [\, F(\xi) - F(\eta)\,] \;+\; \sum_{\zeta \in A} R(\eta,\zeta)\, [\, F(\zeta) - F(\eta)\,] \Big\} G(\eta)$$
$$+\; \sum_{\zeta \in A}\sum_{\xi \in E \setminus A} \pi(\zeta)\, R(\zeta,\xi)\, [\, F(\xi) - F(\zeta)\,]\, G(\zeta)\,.$$

Since F is constant on A, we may add to this expression

$$\sum_{\eta \in A}\sum_{\xi \in A} \pi(\eta)\, R(\eta,\xi)\, [F(\xi) - F(\eta)]\, G(\eta)$$

to obtain that the last displayed expression is equal to $\langle LF, G\rangle_\pi$, which concludes the proof of the first assertion of (A.15). The second statement is obtained following the computation in the reverse order.

Appendix B: Potential theory

In this section, we present general results on the potential theory of continuous-time Markov chains used throughout the article.

Reversible Markov chains can be interpreted in terms of electrical circuits. This description may provide some intuition on the notions introduced below, as Dirichlet form, capacity or equilibrium potential. We refer to the monographs of Doyle and Snell [52] and Gaudillière [66]. The analogy has been extended to the non-reversible context by Balázs and Folly [12].

B.1. The capacity

Fix two non-empty subsets A, B of E such that $A \cap B = \varnothing$. The capacity between A and B, denoted by $\mathrm{cap}(A, B)$, is given by

$$\mathrm{cap}(A, B) := \sum_{\eta \in A} M(\eta) \, \mathbb{P}_\eta[H_B < H_A^+] \,. \tag{B.1}$$

The capacity is monotone in the second coordinate. Let B' be a subset of E such that $A \cap B' = \varnothing$, $B \subset B'$. Since $\mathbb{P}_\eta[H_B < H_A^+] \leq \mathbb{P}_\eta[H_{B'} < H_A^+]$, we have that

$$\mathrm{cap}(A, B) \leq \mathrm{cap}(A, B') \,. \tag{B.2}$$

By (A.4), for any sequence of configurations $\eta_0, \eta_1, \ldots, \eta_n$ such that $p(\eta_i, \eta_{i+1}) > 0$, $0 \leq i < n$,

$$M(\eta_0) \prod_{i=0}^{n-1} p(\eta_i, \eta_{i+1}) \;=\; M(\eta_n) \prod_{i=0}^{n-1} p^*(\eta_{i+1}, \eta_i) \,.$$

In particular, for any $\eta \in A$, $\xi \in B$,

$$M(\eta) \, \mathbb{P}_\eta\big[\, H_B < H_A^+ \,,\, H_B = H_\xi \,\big] \;=\; M(\xi) \, \mathbb{P}_\xi^*\big[\, H_A < H_B^+ \,,\, H_A = H_\eta \,\big] \,.$$

Therefore, since

$$\sum_{\eta \in A} M(\eta) \, \mathbb{P}_\eta\big[\, H_B < H_A^+ \,\big] \;=\; \sum_{\eta \in A} \sum_{\xi \in B} M(\eta) \, \mathbb{P}_\eta\big[\, H_B < H_A^+, H_B = H_\xi \,\big] \,,$$

by (B.1) and the penultimate identity we have that

$$\mathrm{cap}(A, B) \;=\; \sum_{\xi \in B} M(\xi) \, \mathbb{P}_\xi^*[H_A^+ < H_B^+] \;=\; \mathrm{cap}^*(B, A) \,, \tag{B.3}$$

where $\mathrm{cap}^*(A, B)$ represents the capacity between the sets A, B for the adjoint process.

It follows from (B.2) and (B.3) that the capacity is monotone in the first coordinate as well: if A' is a subset of E such that $A \subset A'$, $A' \cap B = \varnothing$,

$$\mathrm{cap}(A, B) \leq \mathrm{cap}(A', B) \,.$$

B.2. A formula for the capacity

Recall the formula (A.7) for the Dirichlet form $D(f)$ of a function $f : E \to \mathbb{R}$. Fix two disjoint subsets A, B of E: $A \cap B = \varnothing$. Denote by $h_{A,B} : E \to \mathbb{R}$ the *equilibrium potential* between A and B. It is the unique solution of the boundary-value elliptic problem

$$\begin{cases} (\,L\,h\,)\,(\eta) \;=\; 0\,, & \eta \in \Omega := (A \cup B)^c\,, \\ h(\eta) \;=\; \chi_A(\eta)\,, & \eta \in A \cup B\,. \end{cases} \tag{B.4}$$

It has a stochastic representation as

$$h_{A,B}(\eta) \;=\; \mathbb{P}_\eta\big[\, H_A < H_B \,\big]\,. \tag{B.5}$$

Since $h_{A,B}$ is harmonic on Ω, it vanishes over B and it is equal to 1 at A,

$$D(h_{A,B}) \;=\; \langle\,(-Lh_{A,B})\,,\, h_{A,B}\,\rangle_\pi \;=\; -\sum_{\eta \in A} \pi(\eta)\,(Lh_{A,B})(\eta)\,. \tag{B.6}$$

By definition of the generator L and since $h_{A,B}$ is equal to 1 on A, the previous expression is equal to

$$D(h_{A,B}) \;=\; \sum_{\eta \in A}\sum_{\xi \in E} \pi(\eta)\,R(\eta,\xi)\,[\,1 - h_{A,B}(\xi)\,]\,.$$

By the representation (B.5) of the equilibrium potential, $1 - h_{A,B}(\xi) = \mathbb{P}_\xi[\,H_B < H_A\,]$. By the strong Markov property at the first jump, for every $\eta \in A$,

$$\mathbb{P}_\eta[\,H_B < H_A^+\,] \;=\; \sum_{\xi \in E} p(\eta,\xi)\,\mathbb{P}_\xi[\,H_B < H_A\,]\,.$$

Hence,

$$D(h_{A,B}) \;=\; \sum_{\eta \in A} \pi(\eta)\,\lambda(\eta)\,\mathbb{P}_\eta[\,H_B < H_A^+\,] \;=\; \mathrm{cap}(A,B)\,. \tag{B.7}$$

The capacity is symmetric: By (B.5), $h_{B,A} = 1 - h_{A,B}$, and, by (A.7), $D(h_{A,B}) = D(1 - h_{A,B})$. Hence,

$$\mathrm{cap}(A,B) \;=\; D(h_{A,B}) \;=\; D(1 - h_{A,B}) \;=\; D(h_{B,A}) \;=\; \mathrm{cap}(B,A)\,. \tag{B.8}$$

B.3. Flows

Denote by $c(\eta,\xi)$ the conductance of the oriented edge (η,ξ), and by $c_s(\eta,\xi)$ its symmetric version:

$$c(\eta,\xi) \;=\; \pi(\eta)\,R(\eta,\xi)\,, \quad c_s(\eta,\xi) \;=\; \frac{1}{2}\,\big\{c(\eta,\xi) + c(\xi,\eta)\big\}\,. \tag{B.9}$$

Note that $c_s(\eta,\xi) = (1/2)\,\pi(\eta)\,\{\,R(\eta,\xi) + R^*(\eta,\xi)\,\}$.

Let $\mathfrak{E}$ be the set of oriented edges defined by

$$\mathfrak{E} \ := \ \{ (\eta, \xi) \in E \times E \ : \ R(\eta, \xi) \,+\, R(\xi, \eta) > 0 \} \ .$$

An anti-symmetric function $\phi : \mathfrak{E} \to \mathbb{R}$ is called a *flow*. The *divergence* of a flow ϕ at $\eta \in E$ is defined as

$$(\mathrm{div}\,\phi)(\eta) \ = \ \sum_{\xi : (\eta, \xi) \in \mathfrak{E}} \phi(\eta, \xi) \ ,$$

while its divergence on a set $A \subset E$ is given by

$$(\mathrm{div}\,\phi)(A) \ = \ \sum_{\eta \in A} (\mathrm{div}\,\phi)(\eta) \ .$$

The flow ϕ is said to be *divergence-free at η* if $(\mathrm{div}\,\phi)(\eta) = 0$.

Denote by $\mathfrak{F}$ the set of flows endowed with the scalar product given by

$$\langle \phi, \psi \rangle \ = \ \frac{1}{2} \sum_{(\eta, \xi) \in \mathfrak{E}} \frac{1}{c_s(\eta, \xi)} \, \phi(\eta, \xi) \, \psi(\eta, \xi) \ , \quad \text{and let} \quad \| \phi \|^2 \ = \ \langle \phi, \phi \rangle \ .$$

Remark B.1. *For a probability measure μ on E, define the flow $\phi_\mu : \mathfrak{E} \to \mathbb{R}$ by $\phi_\mu(\eta, \xi) = \mu(\eta)\,R(\eta, \xi) - \mu(\xi)\,R(\xi, \eta)$. In this set-up, the stationary state corresponds to the probability measure m which turns the flow ϕ_m divergence free at every configuration.*

B.4. The Dirichlet and the Thomson principles

For a function $f : E \to \mathbb{R}$, define the flows Φ_f, Φ_f^* and Ψ_f by

$$\begin{aligned}
\Phi_f(\eta, \xi) \ &= \ f(\eta)\,c(\eta, \xi) \,-\, f(\xi)\,c(\xi, \eta) \ , \\
\Phi_f^*(\eta, \xi) \ &= \ f(\eta)\,c(\xi, \eta) \,-\, f(\xi)\,c(\eta, \xi) \ , \\
\Psi_f(\eta, \xi) \ &= \ c_s(\eta, \xi)\,[\,f(\eta) - f(\xi)\,] \ .
\end{aligned} \tag{B.10}$$

By (A.2) and the fact that the jump rates R^s satisfy the detailed balance conditions (A.3),

$$(\mathrm{div}\,\Phi_f)(\eta) \ = \ -\,\pi(\eta)\,(L^* f)(\eta) \ , \quad (\mathrm{div}\,\Psi_f)(\eta) \ = \ -\,\pi(\eta)\,(L^s f)(\eta) \ . \tag{B.11}$$

It follows from the definition of these flows that for all functions $f : E \to \mathbb{R}$, $g : E \to \mathbb{R}$,

$$\begin{aligned}
\langle \Psi_f, \Phi_g \rangle \ &= \ \langle\,(-L)\,f\,,\,g\,\rangle_\pi \ , \qquad \langle \Psi_f, \Phi_g^* \rangle \ = \ \langle\,(-L^*)\,f\,,\,g\,\rangle_\pi \ , \\
\langle \Psi_f, \Psi_g \rangle \ &= \ \langle\,(-L^s)\,f\,,\,g\,\rangle_\pi \ .
\end{aligned} \tag{B.12}$$

Fix two disjoint subsets A, B of E and two real numbers a, b. Denote by $\mathfrak{C}_{a,b}(A, B)$ the set of functions $f : E \to \mathbb{R}$ which are equal to a on A and b on B:

$$\mathfrak{C}_{a,b}(A, B) \ := \ \big\{ f : E \to \mathbb{R} : f|_A \equiv a, \ f|_B \equiv b \big\} \ .$$

Let $\mathfrak{F}_a(A, B)$ be the set of flows from A to B with strength $a \in \mathbb{R}$:

$$\mathfrak{F}_a(A, B) \;=\; \big\{\, \phi \in \mathfrak{F} \,:\, (\mathrm{div}\,\phi)(A) \,=\, a\,,\, (\mathrm{div}\,\phi)(\eta) = 0\,,\, \eta \in \Omega \big\}\,.$$

In particular, $\mathfrak{F}_1(A, B)$ is the set of *unitary* flows from A to B.

Let $h_{A,B}^*$ be the equilibrium potential corresponding to the adjoint dynamics. It is the solution of the elliptic problem (B.4) with the adjoint generator L^* in place of L. It can be represented through the adjoint chain $\eta^*(t)$ by equation (B.5) with the obvious modifications.

Lemma B.2. *Fix two disjoint and non-empty subsets A, B of E. For every $\alpha \in \mathbb{R}$, $\gamma \in \mathbb{R}$, $f \in \mathfrak{C}_{\alpha,0}(A, B)$, $\phi \in \mathfrak{F}_\gamma(A, B)$,*

$$\langle\, \Phi_f - \phi\,,\, \Psi_{h_{A,B}}\,\rangle \;=\; \alpha\,\mathrm{cap}(A, B) \,-\, \gamma\,.$$

Proof. By definition of the scalar product, the left-hand side is equal to

$$\frac{1}{2}\sum_{(\eta,\xi)\in\mathfrak{E}} \big\{\, \Phi_f(\eta,\xi) - \phi(\eta,\xi)\,\big\}\,\big[\,h_{A,B}(\eta) - h_{A,B}(\xi)\,\big]\,.$$

On the one hand, by definition of Φ_f

$$\frac{1}{2}\sum_{(\eta,\xi)\in\mathfrak{E}} \Phi_f(\eta,\xi)\,\big[\,h_{A,B}(\eta) - h_{A,B}(\xi)\,\big] \;=\; -\sum_{\eta\in E}\pi(\eta)\,f(\eta)\,(Lh_{A,B})(\eta)\,.$$

Since f belongs to $\mathfrak{C}_{\alpha,0}(A, B)$ and $h_{A,B}$ is harmonic on Ω, we may restrict the sum to the set A, and then replace f by $\alpha\,h_{A,B}$ to conclude that the previous sum is equal to

$$-\,\alpha\sum_{\eta\in E}\pi(\eta)\,h_{A,B}(\eta)\,(Lh_{A,B})(\eta) \;=\; \alpha\,D(h_{A,B})\,.$$

By (B.7), this quantity is equal to $\alpha\,\mathrm{cap}(A, B)$.

On the other hand, by definition of the divergence of a flow,

$$-\frac{1}{2}\sum_{(\eta,\xi)\in\mathfrak{E}} \phi(\eta,\xi)\,\big[\,h_{A,B}(\eta) - h_{A,B}(\xi)\,\big] \;=\; -\sum_{\eta\in E}h_{A,B}(\eta)\,(\mathrm{div}\,\phi)(\eta)\,.$$

Since ϕ is divergence-free on Ω and $h_{A,B} = \chi_A$ on $A\cup B$, this expression is equal to

$$-\sum_{\eta\in A}(\mathrm{div}\,\phi)(\eta) \;=\; -\gamma\,.$$

This completes the proof of the lemma. $\qquad\qquad\square$

Theorem B.3 (Dirichlet principle)**.** *For any disjoint and non-empty subsets A, B of E,*

$$\mathrm{cap}(A, B) \;=\; \inf_{f\in\mathfrak{C}_{1,0}(A,B)}\ \inf_{\phi\in\mathfrak{F}_0(A,B)}\ \|\Phi_f - \phi\|^2\,.$$

Furthermore, the unique optimizers of the variational problem are given by

$$f \;=\; \frac{1}{2}(h_{A,B} + h_{A,B}^*) \quad\text{and}\quad \phi \;=\; \frac{1}{2}(\Phi_{h_{A,B}^*} - \Phi_{h_{A,B}}^*)\,.$$

Proof. Fix f in $\mathfrak{C}_{1,0}(A, B)$ and ϕ in $\mathfrak{F}_0(A, B)$. By Lemma B.2 and Schwarz inequality,

$$\mathrm{cap}(A, B)^2 \;=\; \langle\, \Phi_f - \phi\,,\, \Psi_{h_{A,B}}\,\rangle^2 \;\leq\; \|\,\Phi_f - \phi\,\|^2\,\|\,\Psi_{h_{A,B}}\,\|^2 \;.$$

By (B.12) and (B.7),

$$\|\,\Psi_{h_{A,B}}\,\|^2 \;=\; \mathrm{cap}(A, B) \tag{B.13}$$

Hence,

$$\mathrm{cap}(A, B) \;\leq\; \inf_{f \in \mathfrak{C}_{1,0}(A,B)} \inf_{\phi \in \mathfrak{F}_0(A,B)} \|\Phi_f - \phi\|^2 \;.$$

To complete the proof of the theorem, it remains to find f in $\mathfrak{C}_{1,0}(A, B)$ and ϕ in $\mathfrak{F}_0(A, B)$ such that $\|\Phi_f - \phi\|^2 = \mathrm{cap}(A, B)$. We propose $f_0 = (1/2)(h_{A,B} + h^*_{A,B})$ and $\phi_0 = \Phi_{f_0} - \Psi_{h_{A,B}}$.

It is clear that $f_0 = (1/2)(h_{A,B} + h^*_{A,B})$ belongs $\mathfrak{C}_{1,0}(A, B)$. We claim that $\phi_0 = \Phi_{f_0} - \Psi_{h_{A,B}}$ belongs to $\mathfrak{F}_0(A, B)$. By (B.11) and by definition of f_0, the divergence of ϕ_0 at η is equal to

$$-\,\frac{1}{2}\,\pi(\eta) \left\{\, (L^* h_{A,B})(\eta) \;+\; (L^* h^*_{A,B})(\eta) \;-\; (L h_{A,B})(\eta) \;-\; (L^* h_{A,B})(\eta) \,\right\} ,$$

so that

$$(\mathrm{div}\,\phi_0)(\eta) \;=\; \frac{1}{2}\,\pi(\eta) \left\{\, (L h_{A,B})(\eta) \;-\; (L^* h^*_{A,B})(\eta) \,\right\} \;.$$

As $h_{A,B}$ and $h^*_{A,B}$ are harmonic on Ω, ϕ_0 is divergence-free on this set. On the other hand, by (B.7), (B.3) and (B.8),

$$2 \sum_{\eta \in A} (\mathrm{div}\,\phi_0)(\eta) \;=\; D(h_{A,B}) \;-\; D(h^*_{A,B}) \;=\; \mathrm{cap}(A, B) \;-\; \mathrm{cap}^*(A, B) \;=\; 0 \;.$$

This proves that ϕ_0 belongs to $\mathfrak{F}_0(A, B)$.

As f_0, ϕ_0 belong to $\mathfrak{C}_{1,0}(A, B)$, $\mathfrak{F}_0(A, B)$, respectively, and since $\Phi_{f_0} - \phi_0 = \Psi_{h_{A,B}}$, by (B.13),

$$\inf_{f \in \mathfrak{C}_{1,0}(A,B)} \inf_{\phi \in \mathfrak{F}_0(A,B)} \|\Phi_f - \phi\|^2 \;\leq\; \|\Phi_{f_0} - \phi_0\|^2 \;=\; \left\|\Psi_{h_{A,B}}\right\|^2 \;=\; \mathrm{cap}(A, B) \;.$$

This completes the proof of the theorem. $\qquad\qquad\square$

Theorem B.4 (Thomson principle)**.** *For any disjoint and non-empty subsets A, B of E,*

$$\frac{1}{\mathrm{cap}(A, B)} \;=\; \inf_{\psi \in \mathfrak{F}_1(A,B)} \inf_{g \in \mathfrak{C}_{0,0}(A,B)} \|\Phi_g - \psi\|^2 \;.$$

Furthermore, the unique optimizers of the variational problem are given by

$$g \;=\; \frac{1}{2}\,\frac{h^*_{A,B} - h_{A,B}}{\mathrm{cap}(A, B)} \quad and \quad \psi \;=\; \frac{1}{2}\,\frac{\Phi_{h^*_{A,B}} + \Phi^*_{h_{A,B}}}{\mathrm{cap}(A, B)} \;.$$

Proof. Fix a function g in $\mathfrak{C}_{0,0}(A,B)$ and a flow ψ in $\mathfrak{F}_1(A,B)$. By Lemma B.2, Schwarz inequality and (B.13),

$$1 \;=\; \langle\, \Phi_g - \psi\,,\, \Psi_{h_{A,B}}\,\rangle^2 \;\le\; \|\,\Phi_g - \psi\,\|^2\, \mathrm{cap}(A,B)\,,$$

so that

$$\frac{1}{\mathrm{cap}(A,B)} \;\le\; \inf_{\psi\in\mathfrak{F}_1(A,B)}\; \inf_{g\in\mathfrak{C}_{0,0}(A,B)}\; \|\,\Phi_g - \psi\,\|^2\,.$$

To complete the proof of the theorem, it remains to find g in $\mathfrak{C}_{0,0}(A,B)$ and ψ in $\mathfrak{F}_1(A,B)$ such that $\|\,\Phi_g - \psi\,\|^2 = 1/\mathrm{cap}(A,B)$. Let $g_0 = (1/2)\,\{h^*_{A,B} - h_{A,B}\}/\mathrm{cap}(A,B)$, which clearly belongs to $\mathfrak{C}_{0,0}(A,B)$, and let $\psi_0 = \Phi_{g_0} + \Psi_{h_0}$, where $h_0 = h_{A,B}/\mathrm{cap}(A,B)$.

We first claim that ψ_0 belongs to $\mathfrak{F}_1(A,B)$. By (B.11),

$$\begin{aligned}
(\mathrm{div}\,\psi_0)(\eta) \;&=\; -\,\pi(\eta)\,(L^*g_0)(\eta) \;-\; \pi(\eta)\,(L^s h_0)(\eta) \\
&=\; -\,\frac{\pi(\eta)}{2\,\mathrm{cap}(A,B)}\,\big\{(L^*h^*_{A,B})(\eta) \;+\; (L h_{A,B})(\eta)\,\big\}\,.
\end{aligned}$$

This proves that ψ_0 is divergence free on Ω. Moreover, the divergence of ψ_0 on A is given by

$$-\,\frac{1}{2\,\mathrm{cap}(A,B)}\,\Big\{ \sum_{\eta\in A} \pi(\eta)\,(L^*h^*_{A,B})(\eta) \;+\; \sum_{\eta\in A} \pi(\eta)\,(L h_{A,B})(\eta)\,\Big\}\,.$$

By (B.6) and (B.7), this expression is equal to 1, which proves the claim.

By definition of g_0, ψ_0, h_0 and by (B.13),

$$\|\,\Phi_{g_0} - \psi_0\,\|^2 \;=\; \|\,\Psi_{h_0}\,\|^2 \;=\; \frac{1}{\mathrm{cap}(A,B)}\,\cdot$$

This completes the proof of the theorem. $\qquad\square$

Remark B.5. *It follows from the proofs of the theorems that the optimal flow ϕ In Theorem B.3 can be written as*

$$\phi \;=\; \Phi_f - \Psi_h\,,$$

where $f = (1/2)\,\{h^*_{A,B} + h_{A,B}\}$ and $h = h_{A,B}$. Similarly, the optimal flow ψ In Theorem B.4 can be written as

$$\psi \;=\; \Phi_{g_0} + \Psi_{h_0}\,,$$

where $g_0 = (1/2)\,\{h^*_{A,B} - h_{A,B}\}/\mathrm{cap}(A,B)$ and $h_0 = h_{A,B}/\mathrm{cap}(A,B)$.

Theorem B.3 appeared in Gaudillière and Landim [67], and Theorem B.4 is due to Slowik [119]. Similar Dirichlet and Thomson principles are available in the context of diffusions processes, [90, 86].

Remark B.6. *Both theorems require an explicit knowledge of the invariant measure which is not always available in non-reversible dynamics. An important open problem consists therefore to derive formulas for the capacity which do not involve the stationary state.*

Remark B.7. *These variational formulae, expressed as infima, provide simple lower and upper bounds for the capacity. To obtain sharp bounds, good approximations of the harmonic functions are needed to produce test functions and test flows close to the optimal ones. In concrete examples, one of the difficulties is that the test flows constructed are never divergence free, and a correction has to be introduced to remove the divergence of the test flow, [91, 93, 117].*

Remark B.8. *Similar variational formulae for the capacity hold in the context of diffusions [90, 86].*

B.5. Reversible dynamics

In the reversible case, the conductance is symmetric: $c(\eta, \xi) = c(\xi, \eta)$. In particular, all flows Φ_f, Φ_f^*, Ψ_f, introduced in (B.10), coincide, and the optimal flow ϕ of Theorem B.3 vanishes because the equilibrium potentials $h_{A,B}^*$, $h_{A,B}$ are equal. Hence, in the reversible case,

$$\mathrm{cap}(A, B) \;=\; \inf_{f \in \mathfrak{C}_{1,0}(A,B)} \|\Phi_f\|^2 \;=\; \inf_{f \in \mathfrak{C}_{1,0}(A,B)} \langle\, (-L)\, f\,,\, f\, \rangle_\pi \;.$$

where the last identity follows from (B.12). We recover in this way the Dirichlet principle for reversible dynamics:

$$\mathrm{cap}(A, B) \;=\; \inf_{f \in \mathfrak{C}_{1,0}(A,B)} D(f) \tag{B.14}$$

In the Thomson principle, the optimal function g vanishes, and we recover the Thomson principle for reversible dynamics:

$$\frac{1}{\mathrm{cap}(A,B)} \;=\; \inf_{\psi \in \mathfrak{F}_1(A,B)} \|\psi\|^2 \;.$$

In the reversible case, the Thomson principle can also be expressed in terms of functions.

Lemma B.9. *We have that*

$$\frac{1}{\mathrm{cap}(A,B)} \;=\; \inf_f \frac{D(f)}{\left(\, \sum_{\eta \in A} \pi(\eta)\,(L\,f)(\eta)\,\right)^2} \;,$$

where the infimum is carried over all functions $f : E \to \mathbb{R}$ such that $(Lf)(\eta) = 0$ for all $\eta \in E \setminus (A \cup B)$.

Proof. Fix a function $f : E \to \mathbb{R}$ such that $(Lf)(\eta) = 0$ for all $\eta \in E \setminus (A \cup B)$. By Schwarz inequality and equation (A.7) for the Dirichlet form,

$$\left(\frac{1}{2} \sum_{\eta, \xi \in E} \pi(\eta) \, R(\eta, \xi) \, [\, f(\xi) - f(\eta) \,] \, [\, h_{A,B}(\xi) - h_{A,B}(\eta) \,] \right)^2 \leq D(f) \, D(h_{A,B}) \, .$$

As the chain is reversible, the jump rates satisfy the detailed balance conditions (A.3). We may thus rewrite the sum appearing on the left-hand side as

$$- \sum_{\eta, \xi \in E} \pi(\eta) \, R(\eta, \xi) \, [\, f(\xi) - f(\eta) \,] \, h_{A,B}(\eta) \; = \; - \sum_{\eta \in E} \pi(\eta) \, (L \, f)(\eta) \, h_{A,B}(\eta) \, .$$

Since $h_{A,B} = \chi_A$ on $A \cup B$ and $Lf = 0$ on the complement, the previous sum is equal to

$$- \sum_{\eta \in A} \pi(\eta) \, (L \, f)(\eta) \, .$$

We have thus proved that

$$\sup_f \left(\sum_{\eta \in A} \pi(\eta) \, (L \, f)(\eta) \right)^2 \frac{1}{D(f)} \; \leq \; D(h_{A,B}) \, ,$$

where the supremum is carried over all functions f satisfying the assumptions of the lemma. This inequality is actually an identity because the equilibrium potential $h_{A,B}$ belongs to the class of functions considered [it is harmonic on Ω] and

$$\sum_{\eta \in A} \pi(\eta) \, (L \, h_{A,B})(\eta) \; = \; D(h_{A,B}) \, .$$

To complete the proof of the lemma, it remains to recall that $\mathrm{cap}(A, B) = D(h_{A,B})$. □

Remark B.10. *By inserting test functions, the previous lemma provides lower bounds for the capacity between two sets. In practical situations, however, it is almost impossible to find functions which are harmonic at every point of Ω. But it might be possible to find functions which are almost harmonic in the sense that Lf is small. The previous proof applied to any test function yields that for every $\epsilon > 0$,*

$$(1 - \epsilon) \left(\sum_{\eta \in A} \pi(\eta) \, (L \, f)(\eta) \right)^2 - \frac{1}{\epsilon} \left(\sum_{\eta \in \Omega} \pi(\eta) \, | \, (L \, f)(\eta) \, | \right)^2 \; \leq \; D(f) \, D(h_{A,B})$$

where we used Young's inequality $2ab \geq -\epsilon a^2 - \epsilon^{-1} b^2$ and the fact that the absolute value of the harmonic function is bounded by 1. The advantage of this inequality with respect to the Thomson principle lies in the fact that it holds for all functions $f : E \to \mathbb{R}$ and not only for the harmonic ones in Ω. However, the resulting lower bound for the capacity will be sharp only if f is almost harmonic on Ω.

Remark B.11. *The previous remark can be extended to all principles stated in the previous and in the next section. It is this version which is used in concrete examples. We refer to Theorem 5.3 of [117].*

B.6. Dirichlet principle II

We provide in this subsection an alternative variational formula for the capacity in terms of functions only.

Fix two disjoint subsets A, B of E. Let $\mathfrak{F}_0(A, B)^\perp$ be the set of flows in $\mathfrak{F}$ which are orthogonal to all flows in $\mathfrak{F}_0(A, B)$. By [97, Theorem 8.7], for every function f in $\mathfrak{C}_{1,0}(A, B)$,

$$\inf_{\phi \in \mathfrak{F}_0(A,B)} \| \Phi_f - \phi \|^2 \;=\; \sup_{\psi \in \mathfrak{F}_0(A,B)^\perp} \frac{\langle\, \Phi_f\,,\, \psi\,\rangle^2}{\langle\, \psi\,,\, \psi\,\rangle} \;,$$

where the supremum is carried over all $\psi \neq 0$. We may rewrite the right-hand side to obtain that

$$\inf_{\phi \in \mathfrak{F}_0(A,B)} \| \Phi_f - \phi \|^2 \;=\; \sup_{\psi \in \mathfrak{F}_0(A,B)^\perp} \big\{\, 2\,\langle\, \Phi_f\,,\, \psi\,\rangle \;-\; \langle\, \psi\,,\, \psi\,\rangle \,\big\} \;, \qquad \text{(B.15)}$$

which is more convenient.

Let $\mathfrak{C}(A, B)$ be the set of functions $f : E \to \mathbb{R}$ which are constant in A and B:

$$\mathfrak{C}(A, B) \;:=\; \bigcup_{a,b \in \mathbb{R}} \mathfrak{C}_{a,b} \;.$$

Lemma B.12. *We have that*

$$\mathfrak{F}_0(A, B)^\perp \;=\; \big\{\, \Psi_f : f \in \mathfrak{C}(A, B) \,\big\} \;.$$

Proof. Denote by $\mathfrak{A}$ the set on the right-hand side. Its is clear that $\mathfrak{A} \subset \mathfrak{F}_0(A, B)^\perp$. Indeed, fix $\phi \in \mathfrak{F}_0(A, B)$ and f in $\mathfrak{C}_{a,b}$ for some $a, b \in \mathbb{R}$. Then,

$$\langle\, \Psi_f\,,\, \phi\,\rangle \;=\; \frac{1}{2} \sum_{(\eta,\xi) \in \mathfrak{E}} [\, f(\eta) - f(\xi)\,]\, \phi(\eta, \xi) \;=\; \sum_{\eta \in E} f(\eta)\, (\operatorname{div}\phi)(\eta) \;.$$

As f is constant equal to a, b on A, B, respectively, this sum can be written as

$$a \sum_{\eta \in A} (\operatorname{div}\phi)(\eta) \;+\; \sum_{\eta \notin A \cup B} f(\eta)\, (\operatorname{div}\phi)(\eta) \;+\; b \sum_{\eta \in B} (\operatorname{div}\phi)(\eta) \;. \qquad \text{(B.16)}$$

Each of these sums vanish because ϕ belongs to $\mathfrak{F}_0(A, B)$.

It remains to show that $\mathfrak{A}^\perp \subset \mathfrak{F}_0(A, B)$. Let ϕ be a flow in $\mathfrak{A}^\perp$. Then, for all $a, b \in \mathbb{R}$, f in $\mathfrak{C}_{a,b}$,

$$\langle\, \Psi_f\,,\, \phi\,\rangle \;=\; 0 \;.$$

In the first part of the proof, we showed that the left-hand side of this identity is equal to (B.16). Hence, for all $a, b \in \mathbb{R}$ and all $f : E \setminus (A \cup B) \to \mathbb{R}$, (B.16) vanishes. From this we conclude that for all $\xi \notin A \cup B$,

$$\sum_{\eta \in A} (\operatorname{div}\phi)(\eta) \;=\; (\operatorname{div}\phi)(\xi) \;=\; \sum_{\eta \in B} (\operatorname{div}\phi)(\eta) \;=\; 0 \;.$$

This proves that ϕ belongs to $\mathfrak{F}_0(A, B)$ and completes the proof of the lemma. $\qquad\square$

It follows from (B.15), the previous lemma and (B.12) that

$$\inf_{\phi \in \mathfrak{F}_0(A,B)} \|\Phi_f - \phi\|^2 \;=\; \sup_{g \in \mathfrak{C}(A,B)} \left\{ 2 \langle f, Lg \rangle_\pi \,-\, \langle (-L^s g), g \rangle_\pi \right\},$$

where the set $\mathfrak{C}(A,B)$ has been introduced in the statement of Lemma B.12. We replaced g by $-g$ in the previous expression to remove the minus sign in the first term.

The previous argument permitted to formulate in terms of functions a variational formula originally expressed through flows. Since, by (A.6), $\langle L^s g, g \rangle_\pi = \langle Lg, g \rangle_\pi$, in the previous formula we may replace L^s by L. This identity together with Theorem B.3 provides a Dirichlet principle in terms of functions only. This is the content of the next result. In contrast with the one formulate in terms of flows, it involves an inf sup instead of an inf inf which is simpler to estimate.

Theorem B.13. *Let A, B be disjoint, non-empty subsets of E. Then,*

$$\mathrm{cap}(A,B) \;=\; \inf_{f \in \mathfrak{C}_{1,0}(A,B)} \sup_{g \in \mathfrak{C}(A,B)} \left\{ 2 \langle f, Lg \rangle_\pi \,-\, \langle (-L)g, g \rangle_\pi \right\}.$$

Moreover, the optimal function is given by $f = (1/2)\{h_{A,B} + h_{A,B}^*\}$.

Theorem B.13 has been proved by Doyle [51] and, independently, by Gaudillière and Landim [67]. A version in the context of diffusions is due to Pinsky [112, 113].

Remark B.14. *It is also possible to transform the variational problem*

$$\inf_{g \in \mathfrak{C}_{0,0}(A,B)} \|\Phi_g - \psi\|^2$$

into a supremum over flows satisfying certain identities. The resulting variational formula does not seem to be useful.

B.7. Sector condition

Recall from (A.5) that we denote by L^s the symmetric part of the operator L in $L^2(\pi)$: $L^s = (1/2)(L + L^*)$. This operator is self-adjoint in $L^2(\pi)$ and the corresponding Markov chain, denoted by $\eta^s(t)$ is reversible. Moreover, for every function $f : E \to \mathbb{R}$,

$$\langle (-L^s) f, f \rangle_\pi \;=\; \langle (-L) f, f \rangle_\pi \;=\; D(f).$$

Therefore, the Dirichlet form associated to the operator L^s, denoted by $D^s(f)$ and defined by the leftmost term of the previous equation, coincides with the Dirichlet form of the original process.

In particular, if we represent by $\mathrm{cap}^s(A,B)$ the capacity between two disjoint, non-empty subsets A, B with respect to the chain $\eta^s(t)$, by (B.14),

$$\mathrm{cap}^s(A,B) \;=\; \inf_{f \in \mathfrak{C}_{1,0}(A,B)} D_s(f) \;=\; \inf_{f \in \mathfrak{C}_{1,0}(A,B)} D(f).$$

Hence, as $h_{a,B}$ belongs to $\mathfrak{C}_{1,0}(A,B)$, by (B.7) and the previous identity,

$$\operatorname{cap}^s(A,B) \ \le \ \operatorname{cap}(A,B) \ . \tag{B.17}$$

It turns out that a converse inequality holds if the generator satisfies a sector condition. Recall that a generator L satisfies a sector condition with constant C_0 if for every functions f, $g : E \to \mathbb{R}$,

$$\langle\, Lf \,,\, g \,\rangle_\pi^2 \ \le \ C_0 \,\langle\, (-L)\,f \,,\, f \,\rangle_\pi \,\langle\, (-L)\,g \,,\, g \,\rangle_\pi \ .$$

Next result states that the capacity between two sets can be estimated by by the symmetric capacity between these set if the generator satisfies a sector condition

Lemma B.15. *Suppose that the generator L satisfies a sector condition with constant C_0. Then, for every pair of disjoint subsets A, B of E,*

$$\operatorname{cap}(A,B) \ \le \ C_0 \operatorname{cap}^s(A,B) \ .$$

Remark B.16. *By equation* (6.2), *the height of a valley is proportional to the inverse of the capacity. Thus, equation* (B.17) *asserts that the height of a valley in non-reversible dynamics is smaller than the one in the reversible version. Therefore, non-reversible dynamics mix faster than their reversible counterpart.*

Remark B.17. *When the state space E is finite, the generator always satisfies a sector condition (cf. Corollary A.3), but Lemma B.15 holds in the context of countably-infinite state spaces and diffusions.*

B.8. Recurrence

We assume in this section that the set E is countably infinite. A classical problem in the theory of Markov chains is to determine wether a chain is recurrent or not. Potential theory is a powerful tool in this framework.

Here is an open problem, for instance. Consider the random walk in random environment evolving on $\mathbb{Z}^2$ as follows. For each line $l(k) = \{(x,k) : x \in \mathbb{Z}\}$ flip a fair coin. If it comes head, on this line the random walk may only jump to the right, while it may only jump to the left if it comes tail. This represented by drawing an arrow from (x,k) to $(x+1,k)$ for each $x \in \mathbb{Z}$ if the side shown is head, or from (x,k) to $(x-1,k)$ if it is tail. Do the same thing for each column to obtain a graph as in Figure 7.

As illustrated in Figure 7, each point (x,y) in $\mathbb{Z}^2$ is the tail of two arrows. Denote by $\eta(t)$ the random walk on $\mathbb{Z}^2$ which waits a mean-one exponential time ate each site of $\mathbb{Z}^2$ and which jumps with equal probability along one of the two arrows.

It is clear that for almost all environments the random walk is irreducible and that the uniform measure is stationary [because the flow formed by the arrows is divergence-free]. It is an open problem to determine if this random walk is almost-surely recurrent or transient.
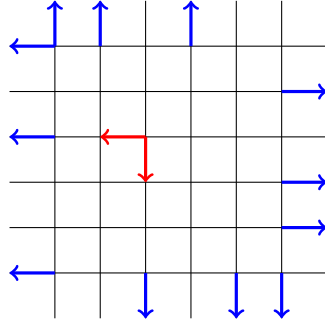
FIG 7. *A random walk in random environment evolving on $\mathbb{Z}^2$. At the tail of the two red arrows, the random walk may only jump, with equal probability, to the left or to the bottom.*

In view of this example, consider a chain $\eta(t)$ defined on a countably infinite space E which is irreducible and assume that there exists a stationary state, denoted by π. Note that π may not be summable, as in the example above. But we assume that π is explicitly known because all estimates below involve π. This is clearly a strong hypothesis and in many cases a stationary state is not known.

Recall that the Markov chain $\eta(t)$ is recurrent if and only if there exist a configuration $\eta \in E$ such that $\mathbb{P}_\eta[H_\eta^+ = \infty] = 0$. There is nothing special about η. If this identity holds for some configuration η, due to the irreducibility, it holds for every. Let $(B_n : n \geq 1)$ be a sequence of *finite* subsets of E containing η and increasing to E, $\eta \in B_n \subset B_{n+1}$, $\cup_n B_n = E$. Then,

$$\mathbb{P}_\eta\big[H_\eta^+ = \infty\big] \;=\; \lim_{n\to\infty} \mathbb{P}_\eta\big[H_{B_n^c} < H_\eta^+\big] \,.$$

By definition (B.1) of the capacity, for any finite set B containing the site 0,

$$\frac{1}{M(\eta)}\,\mathbb{P}_\eta\big[H_{B^c} < H_\eta^+\big] \;=\; \mathrm{cap}(\eta, B^c) \,,$$

where $M(\xi) = \pi(\xi)\,\lambda(\xi)$, $\lambda(\xi)$ being the holding rate at ξ. Hence, the Markov chain $\eta(t)$ is recurrent if and only if there exist a configuration $\eta \in E$ and a sequence of finite subsets B_n containing η and increasing to E such that

$$\lim_{n\to\infty} \mathrm{cap}(\eta, B_n^c) \;=\; 0 \,. \tag{B.18}$$

The proof of the recurrence is thus reduced to the estimation of the capacity between a configuration and the complement of a finite set.

Of course, if condition (B.18) holds for some configuration $\eta \in E$ and for some sequence of finite subsets B_n containing η and increasing to E, it also holds for all configurations $\xi \in E$ and for all sequences of finite subsets C_n containing ξ and increasing to E.

The next two results, taken from [67], follow from the previous observation and the estimate (B.17) and Lemma B.15. Recall from the previous subsection

that $\eta^s(t)$ stands for the reversible version of the process $\eta(t)$ whose generator is given by L^s introduced in (A.5).

Theorem B.18. *Let $\eta(t)$ be a irreducible Markov chain on a countable state space E which admits a stationary measure. The process is transient if the Markov chain $\eta^s(t)$ is.*

Theorem B.19. *Let $\eta(t)$ be a irreducible Markov chain on a countable state space E which admits a stationary measure. The process is recurrent if its generator satisfies a sector condition and if the Markov chain $\eta^s(t)$ is recurrent.*

It follows from these results, cf. [67], that a irreducible Markov chain on a countable state space E which admits a stationary measure is recurrent if the Markov chain $\eta^s(t)$ is recurrent and if

$$\sum_{(\eta,\xi)\in\mathfrak{E}} \frac{c_a(\eta,\xi)^2}{c_s(\eta,\xi)} \; < \; \infty \,,$$

where the symmetric conductance c_s has been introduced in (B.9), and the asymmetric one is given by $c_a(\eta,\xi) = (1/2)\,[\,c(\eta,\xi) - c(\xi,\eta)\,]$.

Benjamini and Hermon [74, 21] used Theorem B.18 to investigate the recurrence of non-backtracking random walks and to show that for every transient, nearest-neighbor Markov chain on a graph, the graph formed by the vertices it visited and edges it crossed is a.s. recurrent for simple random walk.

B.9. Equilibrium measure

Fix two proper, disjoint subsets A, B of E: $A \cap B = \varnothing$, $A \neq \varnothing$, $B \neq \varnothing$. Define the equilibrium measures ν_{AB}, ν_{AB}^* on A as

$$\nu_{AB}(\eta) \;=\; \frac{M(\eta)\,\mathbb{P}_\eta\big[\,H_B < H_A^+\,\big]}{\mathrm{cap}(A,B)}\,, \quad \nu_{AB}^*(\eta) \;=\; \frac{M(\eta)\,\mathbb{P}_\eta^*\big[\,H_B < H_A^+\,\big]}{\mathrm{cap}^*(A,B)} \quad \eta \in A\,.$$

For a probability measure μ on E, denote by $\mathbb{P}_\mu$ the measure on $D([0,\infty),E)$ induced by the Markov chain $\eta(t)$ starting from μ: $\mathbb{P}_\mu = \sum_{\eta\in E} \mu(\eta)\,\mathbb{P}_\eta$. Expectation with respect to $\mathbb{P}_\mu$ is represented by $\mathbb{E}_\mu$.

Proposition B.20. *Fix two proper, disjoint subsets A, B of E. For every function $g : E \to \mathbb{R}$,*

$$\mathbb{E}_{\nu_{AB}^*}\Big[\int_0^{H_B} g(\eta(t))\,dt \Big] \;=\; \frac{\langle\, g\,,\, h_{AB}^*\,\rangle_\pi}{\mathrm{cap}(A,B)} \,. \tag{B.19}$$

Proof. We first claim that the proposition holds for indicator function of a configuration. Fix an arbitrary element ξ of E. If ξ belongs to B the right and the left-hand side of (B.19) vanish. We may therefore assume that ξ does

not belong to B. In this case we may write the expectation appearing in the statement of the lemma as

$$\mathbb{E}_{\nu_{AB}^*}\Big[\sum_{n=0}^{\mathbb{H}_B-1}\frac{e_n}{\lambda(\xi)}\,\chi_{\{\xi\}}(Y_n)\Big]\,,$$

where $(Y_n : n \geq 0)$ is the discrete-time, embedded Markov chain, $(e_n : n \geq 0)$ is a sequence of i.i.d. mean-one, exponential random variables, independent of the jump chain $(Y_n : n \geq 0)$, and $\mathbb{H}_B$ the hitting time of the set B for the discrete time Markov chain Y_n. By the Markov property and by definition of the harmonic measure ν_{AB}^*, this expression is equal to

$$\frac{1}{\lambda(\xi)}\sum_{\eta \in A}\sum_{n \geq 0}\nu_{AB}^*(\eta)\,\mathbb{P}_\eta\big[Y_n = \xi\,,\, n < \mathbb{H}_B\big]$$

$$= \frac{1}{\lambda(\xi)\,\mathrm{cap}^*(A,B)}\sum_{n \geq 0}\sum_{\eta \in A}M(\eta)\,\mathbb{P}_\eta^*\big[H_B < H_A^+\big]\,\mathbb{P}_\eta\big[Y_n = \xi\,,\, n < \mathbb{H}_B\big]\,.$$

We may replace the hitting time and the return time H_B, H_A^+ by the respective times $\mathbb{H}_B$, $\mathbb{H}_A^+$ for the discrete chain. On the other hand, since η and ξ do not belong to B, the event $\{Y_0 = \eta\,,\, Y_n = \xi\,,\, n < \mathbb{H}_B\}$ represents all paths that started from η, reached ξ at time n without passing through B. In particular, by the detailed balanced relations between the process and its adjoint, $M(\eta)\,\mathbb{P}_\eta[Y_n = \xi\,,\, n < \mathbb{H}_B] = M(\xi)\,\mathbb{P}_\xi^*[Y_n = \eta\,,\, n < \mathbb{H}_B]$ and the last sum becomes

$$\frac{M(\xi)}{\lambda(\xi)\,\mathrm{cap}^*(A,B)}\sum_{n \geq 0}\sum_{\eta \in A}\mathbb{P}_\eta^*\big[\mathbb{H}_B < \mathbb{H}_A^+\big]\,\mathbb{P}_\xi^*\big[Y_n = \eta\,,\, n < \mathbb{H}_B\big]$$

$$= \frac{M(\xi)}{\lambda(\xi)\,\mathrm{cap}^*(A,B)}\sum_{n \geq 0}\sum_{\eta \in A}\mathbb{P}_\xi^*\big[Y_n = \eta\,,\, n < \mathbb{H}_B\,,\, \mathbb{H}_B \circ \vartheta_n < \mathbb{H}_A^+ \circ \vartheta_n\big]\,,$$

where we used the Markov property in the last step. In this formula $(\vartheta_k : k \geq 1)$ stands for the group of discrete time shift. Summing over η the sum can be written as

$$\frac{M(\xi)}{\lambda(\xi)\,\mathrm{cap}^*(A,B)}\sum_{n \geq 0}\mathbb{P}_\xi^*\big[Y_n \in A\,,\, n < \mathbb{H}_B\,,\, \mathbb{H}_B \circ \vartheta_n < \mathbb{H}_A^+ \circ \vartheta_n\big]\,.$$

The set inside the probability represents the event that the process Y_k visits A before visiting B and that its last visit to A before reaching B occurs at time n. Hence, since $M(\xi) = \lambda(\xi)\,\pi(\xi)$, since by (B.3), (B.8) $\mathrm{cap}^*(A,B) = \mathrm{cap}(A,B)$ and since g is the indicator function of the configuration ξ, summing over n we get that the previous expression is equal to

$$\frac{1}{\mathrm{cap}^*(A,B)}\,\pi(\xi)\,\mathbb{P}_\xi^*\big[\mathbb{H}_A < \mathbb{H}_B\big] = \frac{\langle g\,,\, h_{AB}^*\rangle_\pi}{\mathrm{cap}(A,B)}\,.$$

By linearity, we get the desired result for all functions $g : E \to \mathbb{R}$. $\qquad\square$

In the particular case where $A = \{\eta\}$ for $\eta \notin B$ we have that

$$\mathbb{E}_\eta \Big[\int_0^{H_B} g(\eta(s)) \, ds \Big] \ = \ \frac{\langle \, g \, , \, h^*_{\{\eta\}B} \, \rangle_\pi}{\mathrm{cap}(\{\eta\}, B)} \tag{B.20}$$

for all functions $g : E \to \mathbb{R}$.

Acknowledgments

References

[1] D. J. Aldous, M. Brown: Inequalities for rare events in time-reversible Markov chains. I. Stochastic inequalities (Seattle, WA, 1991), 1–16, IMS Lecture Notes Monogr. Ser. **22**, Inst. Math. Statist., Hayward, CA, 1992. MR1228050

[2] D. Aldous, J. A. Fill: Reversible Markov chains and random walks on graphs. Available at http://www.stat.berkeley.edu/~aldous/RWG/book.html, (2001).

[3] I. Armendáriz, S. Grosskinsky, M. Loulakis: Zero-range condensation at criticality. Stoch. Proc. Appl. **123**, 3466–3496 (2013). MR3071386

[4] I. Armendáriz, S. Grosskinsky, M. Loulakis: Metastability in a condensing zero-range process in the thermodynamic limit. Probab. Theory Related Fields **169**, 105–175 (2017). MR3704767

[5] I. Armendáriz, M. Loulakis: Thermodynamic limit for the invariant measures in supercritical zero range processes. Probab. Theory Related Fields **145**, 175–188 (2009). MR2520125

[6] I. Armendáriz, M. Loulakis: Conditional distribution of heavy tailed random variables on large deviations of their sum. Stoch. Proc. Appl. **121**, 1138–1147 (2011). MR2775110

[7] S. Arrhenius: On the reaction velocity of the inversion of cane sugar by acids. J. Phys. Chem. **4**, 226 (1889). Partially translated to english in "Selected readings in chemical kinetics", edited by M. Back and K. Laidler, Pergamon press, Oxford 1967.

[8] L. Avena, F. Castell, A. Gaudillière, C. Melot: Approximate and exact solutions of intertwining equations through random spanning forests. arXiv:1702.05992 (2017). MR3866604

[9] L. Avena, F. Castell, A. Gaudillière, C. Melot: Intertwining wavelets or Multiresolution analysis on graphs through random forests. arXiv:1707.04616 (2017). MR3876915

[10] L. Avena, F. Castell, A. Gaudillière, C. Melot: Random Forests and Networks Analysis. arXiv:1711.01635 (2017). MR3876915

[11] L. Avena, A. Gaudillière: Two applications of random spanning forests. J. Theor. Probab., to appear (2018). MR3866604

[12] M. Balázs, A. Folly: An Electric Network for Nonreversible Markov Chains. Amer. Math. Monthly, **123**, 657–682 (2016). MR3539852

[13] J. Beltrán, E. Chavez, C. Landim: From coalescing random walks on a torus to Kingman's coalescent. arXiv:1803.03199 (2018).

[14] J. Beltrán, M. Jara, C. Landim: A martingale problem for an absorbed diffusion: the nucleation phase of condensing zero range processes. Probab. Theory Relat. Fields **169**, 1169–1220 (2017). MR3719065

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

[16] J. Beltrán, C. Landim: Metastability of reversible condensed zero range processes on a finite set. Probab. Theory Relat. Fields **152**, 781–807 (2012). MR2892962

[17] J. Beltrán, C. Landim: Metastability of reversible finite state Markov processes. *Stoch. Proc. Appl.* **121**, 1633–1677 (2011). MR2811018

[18] J. Beltrán, C. Landim: Tunneling of the Kawasaki dynamics at low temperatures in two dimensions. Ann. Inst. Henri Poincaré Probab. Stat. **51**, 59–88 (2015). MR3300964

[19] J. Beltrán, C. Landim: Tunneling and metastability of continuous time Markov chains II. J. Stat. Phys. **149**, 598–618 (2012). MR2998592

[20] J. Beltrán, C. Landim: A Martingale approach to metastability. Probab. Th. Rel. Fields. **161**, 267–307 (2015). MR3304753

[21] I. Benjamini, J. Hermon: Recurrence of Markov chain traces, preprint arXiv:1711.03479v2 (2017).

[22] O. Benois, C. Landim, C. Mourragui: Hitting Times of Rare Events in Markov Chains. J. Stat. Phys. **153**, 967–990 (2013). MR3131505

[23] N. Berglund: Kramers' law: validity, derivations and generalisations. Markov Process. Related Fields, **19**, 459–490 (2013). MR3156961

[24] A. Bianchi, A. Gaudillière: Metastable states, quasi-stationary distributions and soft measures. To appear in Stoch. Proc. Appl. **126**, 1622–1680 (2016). MR3483732

[25] A. Bianchi, S. Dommers, C. Giardinà: Metastability in the reversible inclusion process. Electron. J. Probab. **22**, paper no. 70, (2017). MR3698739

[26] P. Billingsley, *Convergence of Probability Measures, 2nd Edition*. John Wiley & Sons, 1999. MR1700749

[27] M. Blume: Theory of the first-order magnetic phase change in UO_2. Phys. Rev. **141**, 517–524 (1966).

[28] F. Bouchet, J. Reygner: Generalisation of the Eyring-Kramers transition rate formula to irreversible diffusion processes, preprint arxiv:1507.02104 (2015). MR3568024

[29] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein: Metastability in stochastic dynamics of disordered mean field models. Probab. Theory Relat. Fields **119**, 99–161 (2001). MR1813041

[30] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein: Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. J. Eur. Math. Soc. **6**, 399–424 (2004). MR2094397

[31] A. Bovier, F. den Hollander: *Metastability: a potential-theoretic approach.* Grundlehren der mathematischen Wissenschaften **351**, Springer, Berlin, 2015. MR3445787

[32] C. J. Burke, M. Rosenblatt: A Markovian function of a Markov chain. Ann. Math. Statist. **29**, 1112–1122 (1958). MR0101557

[33] M. Cameron, E. Vanden-Eijnden: Flows in Complex Networks: Theory, Algorithms, and Application to Lennard–Jones Cluster Rearrangement. J. Stat. Phys. **156**, 427–454 (2014). MR3217531

[34] J. Cao, P. Chleboun, S. Grosskinsky: Dynamics of condensation in the totally asymmetric inclusion process. J. Stat. Phys. **155**, 523–543 (2014). MR3192172

[35] P. Caputo, H. Lacoin, F. Martinelli, F. Simenhaus and F. L. Toninelli: Polymer dynamics in the depinned phase: metastability with logarithmic barriers. Probab. Theory Related Fields **153**, 587–641 (2012). MR2948687

[36] H. W. Capel: On the possibility of first-order phase transitions in Ising systems of triplet ions with zero-field splitting. Physica **32**, 966–988 (1966); **33**, 295–331 (1967); **37**, 423–441 (1967).

[37] P. Caputo, F. Martinelli and F. L. Toninelli: On the approach to equilibrium for a polymer with adsorption and repulsion. Elect. J. Probab. **13**, 213–258 (2008). MR2386733

[38] M. Cassandro, A. Galves, E. Olivieri, M. E. Vares: Metastable behavior of stochastic dynamics: A pathwise approach. J. Stat. Phys. **35**, 603–634 (1984). MR0749840

[39] Y.-X. Chau, C. Connaughton, S. Grosskinsky: Explosive condensation in symmetric mass transport models. J. Stat. Mechanics: Theory and Experiment, page P11031 (2015).

[40] Y. T. Chen, J. Choi, J. T. Cox: On the convergence of densities of finite voter models to the Wright-Fisher diffusion. Ann. Inst. H. Poincaré Probab. Statist. **52**, 286–322 (2016). MR3449304

[41] P. Chleboun, S. Grosskinsky: Finite size effects and metastability in zero-Range condensation. J. Stat. Phys. **140**, 846–872 (2010). MR2673337

[42] P. Chleboun, S. Grosskinsky: Condensation in stochastic particle systems with stationary product measures. J. Stat. Phys. **154**, 432–465 (2014). MR3162548

[43] P. Chleboun, S. Grosskinsky: A dynamical transition and metastability in a size-dependent zero-range process. Journal of Physics A: Mathematical and Theoretical, **48**, p. 055001, (2015). MR3300272

[44] E. N. M. Cirillo, F. R. Nardi: Relaxation height in energy landscapes: an application to multiple metastable states. *Journal of Statistical Physics* **150**, 1080–1114 (2013). MR3038678

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

[46] E. N. M. Cirillo, F. R. Nardi, C. Spitoni: Sum of exit times in a series of two metastable states. Eur. Phys. J. Spec. Top. **226**, 2421–2438 (2017).

[47] E. N. M. Cirillo, E. Olivieri: Metastability and nucleation for the Blume-Capel model. Different mechanisms of transition. J. Stat. Phys. **83**, 473–554 (1996). MR1386350

[48] C. Cooper, A. Frieze, T. Radzik: Multiple random walks in random regular graphs. SIAM J. Discrete Math. **23**, 1738–1761 (2009). MR2570201

[49] A. Cortines, J. Gold, O. Louidor: Dynamical freezing in a spin glass system with logarithmic correlations. Electron. J. Probab. **23**, paper 59 (2018). MR3814253

[50] J. T. Cox: Coalescing random walks and voter model consensus times on the torus in Z^d. Ann. Probab. **17**, 1333–1366 (1989). MR1048930

[51] P. Doyle: Energy for Markov chains. Unpublished manuscript available at http://www.math.dartmouth.edu/~doyle (1994).

[52] P. G. Doyle, J. L. Snell: *Random walks and electric networks*, volume 22 of Carus Mathematical Monographs. Mathematical Association of America, Washington, DC, 1984. http://arxiv.org/abs/math/0001057. MR0920811

[53] R. Durrett: Some features of the spread of epidemics and information on a random graph. Proc. Nat. Acad. Sci. USA **107** 4491–4498 (2010).

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

[55] S. N. Ethier, Th. G. Kurtz: *Markov processes. Characterization and convergence.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. MR0838085

[56] M. R. Evans: Phase transitions in one-dimensional nonequilibrium systems. Braz. J. Phys. **30**, 42–57 (2000).

[57] L. C. Evans, P. R. Tabrizian: Asymptotic for scaled Kramers-Smoluchowski equations. SIAM J. Math. Anal. **48**, 2944–2961 (2016). MR3542005

[58] H. Eyring: The activated complex in chemical reactions. J. Chem. Phys. **3**, 107–115 (1935).

[59] L. Fajfrová, T. Gobron, E. Saada: Invariant measures of mass migration processes. Electron. J. Probab. **21**, article 60, (2016). MR3563888

[60] R. Fernandez, F. Manzo, F. Nardi, E. Scoppola, J. Sohier: Conditioned, quasi-stationary, restricted measures and metastability. Ann. Appl. Probab. (2015). MR3476624

[61] R. Fernandez, F. Manzo, F. Nardi, E. Scoppola: Asymptotically exponential hitting times and metastability: a pathwise approach without reversibility. Electron. J. Probab. (2015). MR3425542

[62] L. R. G. Fontes, P. H. S. Lima: Convergence of symmetric trap models in the hypercube. *New Trends in Mathematical Physics*, 285–297, ed. V. Sidoravičius, Springer, Netherlands (2009).

[63] L. R. G. Fontes, P. Mathieu: K-processes, scaling limit and aging for the trap model in the complete graph. Ann. Probab. **36**, 1322–1358 (2008). MR2435851

[64] M. I. Freidlin, A. D. Wentzell: Random perturbations of dynamical systems. Second edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **260**. Springer-Verlag, New York, 1998. MR1652127

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

[66] A. Gaudillière: Condenser physics applied to Markov chains: A brief introduction to potential theory. arXiv:0901.3053 (2009).

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

[68] C. Godrèche: Dynamics of condensation in zero-range processes. J. Phys. A: Math. Gen. **36**, 6313–6328 (2003). MR1987434

[69] C. Godrèche and J. M. Luck: Condensation in the inhomogeneous zero-range process: an interplay between interaction and diffusion disorder. J. Stat. Mechanics: Theory and Experiment, P12013 (2012). MR3036143

[70] B. Gois and C. Landim: Zero-temperature limit of the Kawasaki dynamics for the Ising lattice gas in a large two-dimensional torus. Ann. Probab. **43**, 2151–2203 (2015). MR3353824

[71] S. Grosskinsky, F. Redig, K. Vafayi: Condensation in the inclusion process and related models. J. Stat. Phys. **142**, 952–974 (2011). MR2781716

[72] S. Grosskinsky, F. Redig, K. Vafayi: Dynamics of condensation in the symmetric inclusion process **18**, article no. 66 (2013). MR3078025

[73] S. Grosskinsky, G. M. Schütz, H. Spohn: Condensation in the zero range process: stationary and dynamical properties. J. Statist. Phys. **113**, 389–410 (2003). MR2013129

[74] J. Hermon: Reversibility of the non-backtracking random walk, preprint arXiv:1707.01601v1 (2017).

[75] J. van't Hoff: *Studies in Chemical Dynamics*. Chemical publishing Company, 1896.

[76] J. Hood: Phil. Mag. **6**, 371 (1878) and **20**, 323 (1885).

[77] P. Imkeller, I. Pavlyukevich: Metastable behaviour of small noise Lévy-driven diffusions. ESAIM PS 12, 412–437 (2008). MR2437717

[78] M. Jara, C. Landim, A. Teixeira: Quenched scaling limits of trap models. Ann. Probab. **39**, 176–223 (2011). MR2778800

[79] M. Jara, C. Landim, A. Teixeira: Universality of trap models in the ergodic time scale. Ann. Probab. **42**, 2497–2557 (2014). MR3265173

[80] I. Jeon, P. March, B. Pittel: Size of the largest cluster under zero-range invariant measures. Ann. Probab. **28**, 1162–1194 (2000). MR1797308

[81] C. Kipnis, C. M. Newman: The metastable behavior of infrequently observed, weakly random, one dimensional diffusion processes. SIAM J. Appl. Math. **45**, 972–982 (1985). MR0813459

[82] T. Komorowski, C. Landim, S. Olla: *Fluctuations in Markov Processes, Time Symmetry and Martingale Approximation.* Die Grundlehren der mathematischen Wissenschaften **345**, Springer-Verlag, Berlin 2012. MR2952852

[83] H. A. Kramers: Brownian motion in a field of force and the diffusion model of chemical reactions. Physica **7**, 284–304 (1940). MR0002962

[84] H. Lacoin, A. Teixeira: A mathematical perspective on metastable wetting. Electron. J. Probab. **20**, paper no. 17 (2015). MR3317159

[85] 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). MR3215575

[86] C. Landim: Variational formulae for the capacity induced by second-order elliptic differential operators. Proceedings of the International Congress of Mathematicians 2018 (ICM 2018).

[87] C. Landim, P. Lemire: Metastability of the two-dimensional Blume-Capel model with zero chemical potential and small magnetic field. J. Stat. Phys. **164**, 346–376 (2016). MR3513256

[88] C. Landim, P. Lemire, M. Mourragui: Metastability of the two-dimensional Blume-Capel model with zero chemical potential and small magnetic field on a large torus. To appear in J. Stat. Phys. (2019). MR3513256

[89] 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). MR3858923

[90] C. Landim, M. Mariani, I. Seo: A Dirichlet and a Thomson principle for non-selfadjoint elliptic operators, metastability in non-reversible diffusion processes. To appear in Arch. Ration. Mech. Anal., arXiv:1701.00985 (2017). MR3900816

[91] C. Landim, R. Misturini, K. Tsunoda: Metastability of reversible random walks in potential field. J. Stat. Phys. **160**, 1449–1482 (2015). MR3382755

[92] C. Landim, I. Seo: Metastability of non-reversible mean-field Potts model with three spins. J. Stat. Phys. **165**, 693–726 (2016). MR3568163

[93] 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). MR3745152

[94] C. Landim, I. Seo: Metastability of one-dimensional, non-reversible diffusions with periodic boundary conditions. To appear in Ann. Inst. H. Poincaré, Probab. Statist., arXiv:1710.06672 (2017). MR3568163

[95] 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). MR3536686

[96] D. A. Levin, Y. Peres, E. L. Wilmer: *Markov Chains and Mixing Times.* Amer. Math. Soc., Providence, 2009. MR2466937

[97] P. D. Lax: *Functional analysis.* Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2002. MR1892228

[98] J. Lu, E. Vanden-Eijnden: Exact dynamical coarse-graining without time-scale separation. J. Chem. Phys. **141**, 044109 (2014).

[99] F. Manzo, F. Nardi, E. Olivieri, E. Scoppola: On the essential features of metastability: tunnelling time and critical configurations. J. Stat. Phys. **115**, 591–642 (2004). MR2070109

[100] F. Manzo, E. Olivieri: Dynamical Blume-Capel model: competing metastable states at infinite volume. J. Stat. Phys. **104**, 1029–1090 (2001). MR1858997

[101] P. Metzner, Ch. Schuette, E. Vanden-Eijnden: Transition path theory for Markov jump processes. SIAM Multiscale Model. Simul. **7**, 1192–1219 (2009). MR2480117

[102] R. Misturini: Evolution of the ABC model among the segregated configurations in the zero-temperature limit. Ann. Inst. H. Poincaré Probab. Statist. **52**, 669–702 (2016). MR3498005

[103] F. R. Nardi, A. Zocca: Tunneling behavior of Ising and Potts models in the low-temperature regime. arXiv:1708.09677 (2017).

[104] E. J. Neves, R. H. Schonmann: Critical droplets and metastability for a Glauber dynamics at very low temperatures. Comm. Math. Phys. **137**, 209–230 (1991). MR1101685

[105] E. J. Neves, R. H. Schonmann: Behavior of droplets for a class of Glauber dynamics at very low temperature. Probab. Theory Related Fields **91**, 331–354 (1992). MR1151800

[106] C. Oh, F. Rezakhanlou: Metastability of zero range processes via Poisson equations, preprint (2019).

[107] R. I. Oliveira: On the coalescence time of reversible random walks. Trans. Am. Math. Soc. **364**, 2109–2128 (2012). MR2869200

[108] R. I. Oliveira: Mean field conditions for coalescing random walks. Ann. Probab. **41**, 3420–3461 (2013). MR3127887

[109] E. Olivieri, E. Scoppola: Markov Chains with Exponentially Small Transition Probabilities: First Exit Problem from a General Domain. I. The Reversible Case J. Stat. Phys. **79**, 613–647 (1995). MR1327899

[110] E. Olivieri, E. Scoppola: Markov Chains with Exponentially Small Transition Probabilities: First Exit Problem from a General Domain. II. The General Case J. Stat. Phys. **84**, 987–1041 (1996). MR1412076

[111] E. Olivieri and M. E. Vares: *Large deviations and metastability.* Encyclopedia of Mathematics and its Applications, vol. 100. Cambridge University Press, Cambridge, 2005. MR2123364

[112] R. G. Pinsky: A generalized Dirichlet principle for second order non-selfadjoint elliptic operators. SIAM J. Math. Anal. **19**, 204–213 (1988). MR0924555

[113] R. G. Pinsky: A minimax variational formula giving necessary and sufficient conditions for recurrence or transience of multidimensional diffusion processes. Ann. Probab. **16**, 662–671 (1988). MR0929069

[114] F. Rezakhanlou, I. Seo: Scaling limit of metastable diffusions, preprint (2018).

[115] L. C. G. Rogers, D. Williams: *Diffusions, Markov Processes, and Martingales: Volume 1, Foundations.* Cambridge University Press, 1994. MR1780932

[116] E. Scoppola: Renormalization group for Markov chains and application to metastability. J. Stat. Phys. **73**, 83–121 (1993). MR1247859

[117] I. Seo: Condensation of non-reversible zero-range processes, arXiv:1801.05934 (2018). MR3922538

[118] I. Seo, P. R. Tabrizian: Asymptotics for scaled Kramers-Smoluchowski equations in several dimensions with general potentials, preprint (2017).

[119] M. Slowik: A note on variational representations of capacities for reversible and nonreversible Markov chains. Unpublished, Technische Universität Berlin, 2012.

[120] D. W. Stroock, S. R. S. Varadhan: *Multidimensional diffusion processes.* Reprint of the 1997 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2006. MR2190038

[121] M. Sugiura: Metastable behaviors of diffusion processes with small parameter. J. Math. Soc. Japan **47**, 755–788 (1995). MR1348758

[122] M. Sugiura: Asymptotic behaviors on the small parameter exit problems and the singularly perturbation problems. Ryukyu Math. J. **14**, 79–118 (2001). MR1897280

[123] B. Waclaw, M. R. Evans: Explosive Condensation in a Mass Transport Model. Phys. Rev. Lett. **108**, 070601 (2012).