A MARTINGALE APPROACH TO METASTABILITY

J. BELTRÁN, C. LANDIM

ABSTRACT. We presented in [1, 5] an approach to derive the metastable behavior of continuous-time Markov chains. We assumed in these articles that the Markov chains visit points in the time scale in which it jumps among the metastable sets. We replace this condition here by assumptions on the mixing times and on the relaxation times of the chains reflected at the boundary of the metastable sets.

1. Introduction

Cassandro et al. proposed in a seminal paper [13] a general method to derive the metastable behavior of continuous-time Markov chains with exponentially small jump rates, called the pathwise approach. In many different contexts these ideas permitted to prove that the exit time from a metastable set has an asymptotic exponential law; to provide estimates for the expectations of the exit times; to describe the typical escape trajectory from a metastable set; to compute the distribution of the exit (saddle) points from a metastable set; and to prove the convergence of the finite-dimensional distributions of the order parameter, the macroscopic variable which characterizes the state of the process, to the finite-dimensional distributions of a finite-state Markov chain. This approach has known a great success, and it is impossible to review here the main results. We refer to [23] for a recent account of this theory.

In [9, 10], Bovier et al. proposed a new approach to prove the metastable behavior of continuous-time Markov chains, known as the potential theoretic approach. Motivated by the dynamics of mean field spin systems, the authors created tools, based on the potential theory of reversible Markov processes, to compute the expectation of the exit time from a metastable set and to prove that these exit times are asymptotically exponential. They also expressed the expectation of the exit time from a metastable set and the jump probabilities among the metastable sets in terms of eigenvalues and right-eigenvectors of the generator of the Markov chain.

Compared to the pathwise approach, the potential theoretic approach does not attempt to describe the typical exit path from a metastable set, but provides precise asymptotic formulas for the expectation of the exit time from a metastable set. This accuracy, not reached by the pathwise approach, whose estimates admit exponential errors in the parameter, permits to encompass in the theory dynamics which present logarithmic energy or entropy barriers such as [12, 2, 11]. Moreover, in the case of a transition from a metastable set to a stable set, it characterizes the asymptotic dynamics: the process remains at the metastable set an exponential time whose mean has been estimated sharply and then it jumps to the stable set.

Key words and phrases. Metastability, Mixing times, Markov processes.
As the pathwise approach, the potential theoretic approach has been successfully applied to a great number of models. We refer to the recently published paper [6] for references.

Inspired by the evolution of sticky zero-range processes [2, 21], dynamics which have a finite number of stable sets with logarithmic energy barriers, we proposed in [1, 5] a third approach to metastability, now called the martingale approach. This method was successfully applied to derive the asymptotic behavior of the condensate in sticky zero-range processes [2, 21], to prove that in the ergodic time scale random walks among random traps [17, 18] converge to $K$-processes, and to show that the evolution among the ground states of the Kawasaki dynamics for the two dimensional Ising lattice gas [4, 16] on a large torus converges to a Brownian motion as the temperature vanishes.

To depict the asymptotic dynamics of the order parameter, one has to compute the expectation of the holding times of each metastable set and the jump probabilities amid the metastable sets. The potential theoretic approach permits to compute the expectations of the holding times and yields a formula for the jump probabilities in terms of eigenvectors of the generator. This latter formula, although interesting from the theoretical point of view, since it establishes a link between the spectral properties of the generator and the metastable behavior of the process, is of little practical use because one is usually unable to compute the eigenvectors of the generator.

The martingale approach replaces the formula of the jump probabilities written through eigenvectors of the generator by one, [1, Remark 2.9 and Lemma 6.8], expressed only in terms of the capacities, capacities which can be estimated using the Dirichlet and the Thomson variational principles. We have, therefore, a precise description of the asymptotic dynamics of the order parameter: a sharp estimate of the holding times at each metastable set from the potential theoretical approach, and an explicit expression for the jump probabilities among the metastable sets from the aforementioned formula.

This informal description of the asymptotic dynamics of the order parameter among the metastable sets has been converted in [1, 5] into a theorem which asserts that the order parameter converges to a Markov chain in a topology introduced in [18], weaker than the Skorohod one. The proof of this result relies on three hypotheses, formulated in terms of the stationary measure and of the capacities between sets, and it uses the martingale characterization of a Markovian dynamics and the notion of the trace of a Markov process on a subset of the configuration space.

In the martingale approach, the potential theory tools developed by Bovier et al. [9, 10] to prove the metastability of Markov chains can be very useful in some models [2, 21] or not needed at all, as in [17, 18]. In these latter dynamics, the asymptotic jump probabilities among the metastable sets, which, as we said, can be expressed through capacities, are estimated by other means without reference to potential theory.

The proof of the convergence of the order parameter to a Markov chain presented in [1, 5] requires that in each metastable set the time it takes for the process to visit a representative configuration of the metastable set is small compared to the time the process stays in the metastable set. We introduced in [1] a condition, expressed in terms of capacities, which guarantees that a representative point of
the metastable set is visited before the process reaches another metastable set. This quite strong assumption, fulfilled by a large class of dynamics, fails in some cases, as in polymer models in the depinned phase [12, 11] or in the dog graph [24]. The main goal of this article is to weaken this assumption.

More recently, Bianchi and Gaudillièere [7] proposed still another approach based on the fact that the exit time from a set starting from the quasi-stationary measure associated to this set is an exponential random variable. The proof that the exit time from a metastable set is asymptotically exponential is thus reduced to the proof that the state of the process gets close to the quasi-stationary state before the process leaves the metastable set. To derive this property the authors obtained estimates on the mixing time towards the quasi-stationary state and on the asymptotic exit distribution with errors expressed in terms of the ratio between the spectral radius of the generator of the process killed when it leaves the metastable set and the spectral gap of the process reflected at the boundary of the metastable set, a ratio which has to be small if a metastable behavior is expected. They also introduced $(\kappa, \lambda)$-capacities, an object which plays an important role in this article.

After these historical remarks, we present the main results of this article. Consider a sequence of continuous-time Markov chains $\eta^N(t)$. To describe the asymptotic evolution of the dynamics among the metastable sets, let $X^N_t$ be the functional of the process which indicates the current metastable set visited:

$$X^N_t = \sum_{x=1}^{\kappa} x \{ \eta^N(t) \in \mathcal{E}^x_N \}.$$ 

In this formula, $\kappa$ represents the number of metastable sets and $\mathcal{E}^x_N$, $1 \leq x \leq \kappa$, the metastable sets. The non-Markovian dynamics $X^N_t$ is called the order process or the the order in short.

The main result of [1, 5] states that under certain conditions, which can be expressed only in terms of the stationary measure and of the capacities between the metastable sets, the order converges in some time scale and in some topology to a Markov process on $S = \{1, \ldots, \kappa\}$.

The main drawback of the method [1, 5] is that it requires the process to visit points. More precisely, we needed to assume that each metastable set $\mathcal{E}^x_N$ contains a configuration $\xi^x_N$ which, once the process enters $\mathcal{E}^x_N$, is visited before the process reaches another metastable set:

$$\lim_{N \to \infty} \sup_{\eta \in \mathcal{E}^x_N} \mathbb{P}_{\eta}[H^{\xi^x_N}_{\mathcal{E}^x_N} < H_{\xi^x_N}] = 0$$ (1.1)

for all $x \in S$. Here, $H_A$, $A \subset E_N$, stands for the hitting time of $A$, $\mathcal{E}^x_N = \cup_{y \neq x} \mathcal{E}^y_N$, and $\mathbb{P}_{\eta}$ represents the distribution of the process $\eta^N(t)$ starting from the configuration $\eta$. The configuration $\xi^x_N$ is by no means special. It is shown in [1] that if this property holds for one configuration $\xi$ in $\mathcal{E}^x_N$, it holds for any configuration in $\mathcal{E}^x_N$.

Property (1.1) is fulfilled by some dynamics, as sticky zero-range processes [2, 21], trap models [17, 18] or Markov processes on finite sets [3, 4], but it is clearly not fulfilled in general.

The purpose of this paper is to replace condition (1.1) by assumptions on the relaxation time of the process reflected at the boundary of a metastable set. We propose two different set of hypotheses. The first set essentially requires only the
spectral gap of the process to be much smaller than the spectral gaps of the reflected processes on each metastable set, and the average jump rates among the metastable sets to converge when properly renormalized. Under these conditions, Theorem 2.2 states that the finite-dimensional distributions of the order process converge to the finite-dimensional distributions of a finite state Markov chain, provided the initial distribution is not too far from the equilibrium measure.

On the other hand, if one is able to show that the mixing times of the reflected processes on each metastable set are much smaller than the relaxation time of the process, Theorem 2.4 and Lemma 2.5 affirm that the order process converges to a finite state Markov chain. Hence, the condition that the process visits points is replaced in this article by estimates on the mixing times of the reflected processes.

In Section 8, we apply these results to two models. We show that the polymer in the depinned phase considered by Caputo et al. in [12, 11] satisfy the first set of conditions and that the dog graph introduced by Diaconis and Saloff-Coste [24] satisfy the second set of assumptions. H. Lacoin and A. Teixeira [20] are presently working on another polymer model in which the second set of conditions can be verified.

2. Notation and results

Fix a sequence \((E_N : N \geq 1)\) of countable state spaces. The elements of \(E_N\) are denoted by the Greek letters \(\eta, \xi\). For each \(N \geq 1\) consider matrix \(R_N : E_N \times E_N \to \mathbb{R}\) such that \(R_N(\eta, \xi) \geq 0, \eta \neq \xi, -\infty < R_N(\eta, \eta) < 0, \sum_{\xi} R_N(\eta, \xi) = 0, \eta \in E_N\). Denote by \(\{\eta^N(t) : t \geq 0\}\) the right-continuous, continuous-time strong Markov process on \(E_N\) whose generator \(L_N\) is given by

\[
(L_N f)(\eta) = \sum_{\xi \in E_N} R_N(\eta, \xi) \left\{ f(\xi) - f(\eta) \right\},
\]

for bounded functions \(f : E_N \to \mathbb{R}\). We assume that \(\eta^N(t)\) is positive-recurrent and reversible. Denote by \(\pi = \pi_N\) the unique invariant probability measure, by \(\lambda_N(\eta), \eta \in E_N\), the holding rates, \(\lambda_N(\eta) = \sum_{\xi \neq \eta} R_N(\eta, \xi)\), and by \(p_N(\eta, \xi), \eta, \xi \in E_N\), the jump probabilities: \(p_N(\eta, \xi) = \lambda_N(\eta)^{-1} R_N(\eta, \xi)\) for \(\eta \neq \xi\), and \(p_N(\eta, \eta) = 0\) for \(\eta \in E_N\). We assume that \(p_N(\eta, \xi)\) are the transition probabilities of a positive-recurrent discrete-time Markov chain. In particular the measure \(M_N(\eta) := \pi_N(\eta)\lambda_N(\eta)\) is finite.

Throughout this article we omit the index \(N\) as much as possible. We write, for instance, \(\eta(t), \pi\) for \(\eta^N(t), \pi_N\), respectively. Denote by \(D(\mathbb{R}^+, E_N)\) the space of right-continuous trajectories with left limits endowed with the Skorohod topology. Let \(\mathbb{P}_\eta = \mathbb{P}_{\eta}^N, \eta \in E_N\), be the probability measure on \(D(\mathbb{R}^+, E_N)\) induced by the Markov process \(\{\eta(t) : t \geq 0\}\) starting from \(\eta\). Expectation with respect to \(\mathbb{P}_\eta\) is denoted by \(\mathbb{E}_\eta\).

For a subset \(A\) of \(E_N\), denote by \(H_A\) the hitting time of \(A\) and by \(H_A^+\) the return time to \(A\):

\[
H_A^+ = \inf\{t > 0 : \eta(t) \in A, \eta(s) \neq \eta(0)\ \text{for some} \ 0 < s < t\},
\]

\[
H_A := \inf\{t > 0 : \eta(t) \in A\},
\]

with the convention that \(H_A = \infty, H_A^+ = \infty\) if \(\eta(s) \notin A\) for all \(s > 0\). We sometimes write \(H(A)\) for \(H_A\). Denote by \(\text{cap}_N(A, B)\) the capacity between two
disjoint subsets $A, B$ of $E_N$:

$$\text{cap}_N(A, B) = \sum_{\eta \in A} \pi(\eta) \lambda(\eta) \mathbb{P}_\eta[H_B < H_A^+].$$

Denote by $L^2(\pi)$ the space of square summable functions $f : E_N \to \mathbb{R}$ endowed with the scalar product $\langle f, g \rangle_\pi = \sum_{\eta \in E_N} \pi(\eta) f(\eta) g(\eta)$. Let $g = g_N$ be the spectral gap of the generator $L_N$:

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

where the infimum is carried over all functions $f$ in $L^2(\pi)$ which are orthogonal to the constants: $\langle f, 1 \rangle = 0$.

Fix a finite number of disjoint subsets $E^1_N, \ldots, E^\kappa_N$, $\kappa \geq 2$, of $E_N$: $E^i_N \cap E^j_N = \emptyset$, $x \neq y$. The sets $E^i_N$ have to be interpreted as wells for the Markov dynamics $\eta(t)$. Let $E_N = \cup_{x \in S} E^x_N$ and let $\Delta_N = E_N \setminus E_N$ so that

$$E_N = E^1_N \cup \cdots \cup E^\kappa_N \cup \Delta_N. \quad (2.3)$$

In contrast with the wells $E^i_N$, $\Delta_N$ is a set of small measure which separates the wells.

**A. Trace process.** Denote by $\{\eta^N(t) : t \geq 0\}$ the $E_N$-valued Markov process obtained as the trace of $\{\eta^N(t) : t \geq 0\}$ on $E_N$. We refer to [1, Section 6.1] for a precise definition. The rate at which the trace process jumps from $\eta$ to $\xi \in E_N$ is denoted by $R^\xi(\eta, \xi)$ and its generator by $L^\xi$:

$$(L^\xi f)(\eta) = \sum_{\xi \in E_N} R^\xi(\eta, \xi) \left\{ f(\xi) - f(\eta) \right\}, \quad \eta \in E_N.$$ 

By [1, Proposition 6.3], the probability measure $\pi$ conditioned to $E_N$, $\pi^\xi(\eta) = \pi(\eta)/\pi(E_N)$, is reversible for the trace process.

Let $\mathbb{P}_\eta^\xi$, $\eta \in E_N$, be the probability measure on $D(\mathbb{R}_+, E_N)$ induced by the trace process $\{\eta^\xi(t) : t \geq 0\}$ starting from $\eta$. Expectation with respect to $\mathbb{P}_\eta^\xi$ is denoted by $\mathbb{E}_\eta^\xi$. Denote by $g_\xi$ the spectral gap of the trace process:

$$g_\xi = \inf_f \frac{\langle (-L^\xi)f, f \rangle_{\pi^\xi}}{\langle f, f \rangle_{\pi^\xi}},$$

where the infimum is carried over all functions $f$ in $L^2(\pi^\xi)$ which are orthogonal to the constants: $\langle f, 1 \rangle_{\pi^\xi} = 0$.

Proposition [21] presents an estimate of the spectral gap of the trace process in terms of the spectral gap of the original process.

**Proposition 2.1.** Let $f$ be an eigenfunction associated to $g$ such that $E(\pi)[f^2] = 1$, $E(\pi)[f] = 0$. Then,

$$g_\xi \left\{ 1 - \frac{1}{\pi(E_N)} E(\pi)[f^2 1_{E_N}] \right\} \leq g \leq g_\xi.$$

In the examples we have in mind $\pi(E_N)$ converges to 1. In particular, if we show that an eigenfunction associated to $g$ is bounded, $g_\xi/g$ converges to 1. We provide in Lemma [6.1] an upper bound for $g_\xi$ in terms of capacities.
Denote by $\Psi_N : E_N \mapsto S = \{1, \ldots, \kappa\}$, the projection given by

$$\Psi_N(\eta) = \sum_{x=1}^\kappa x 1\{\eta \in E_N^x\}.$$ 

and by $\{X_t^N : t \geq 0\}$ the stochastic process on $S$ defined by $X_t^N = \Psi_N(\eta^N(t))$. Clearly, besides trivial cases, $\{X_t^N : t \geq 0\}$ is not Markovian. We refer to $X_t^N$ as the order process or order for short.

**B. Reflected process.** Denote by $\{\eta^{x,x}(t) : t \geq 0\}$, $1 \leq x \leq \kappa$, the Markov process $\eta(t)$ reflected at $E_N^x$. This is the process obtained from the Markov process $\eta(t)$ by forbidding all jumps from $\eta$ to $\xi$ if $\eta$ or $\xi$ do not belong to $E_N^x$. The generator $L_{\eta,x}$ of this Markov process is given by

$$(L_{\eta,x}f)(\eta) = \sum_{\xi \in E_N^x} R_N(\eta, \xi) \{f(\xi) - f(\eta)\}, \quad \eta \in E_N^x.$$ 

Assume that the reflected process $\eta^{x,x}(t)$ is irreducible for each $1 \leq x \leq \kappa$. It is easy to show that the conditioned probability measure $\pi_x$ defined by

$$\pi_x(\eta) = \frac{\pi(\eta)}{\pi(E_N^x)}, \quad \eta \in E_N^x,$$ 

is reversible for the reflected process. Let $g_{\eta,x}$ be the spectral gap of the reflected process:

$$g_{\eta,x} = \inf_f \frac{\langle (-L_{\eta,x})f, f \rangle_{\pi_x}}{\langle f, f \rangle_{\pi_x}},$$

where the infimum is carried over all functions $f$ in $L^2(\pi_x)$ which are orthogonal to the constants: $\langle f, 1 \rangle_{\pi_x} = 0$.

**C. Enlarged process.** Consider a irreducible, positive recurrent Markov process $\xi(t)$ on a countable set $E$ which jumps from a state $\eta$ to a state $\xi$ at rate $R(\eta, \xi)$. Denote by $\pi$ the unique stationary state of the process. Let $E^*$ be a copy of $E$ and denote by $\eta^* \in E^*$ the copy of $\eta \in E$. Following \[, for $\gamma > 0$ denote by $\gamma^{\gamma}(t)$ the Markov process on $E \cup E^*$ whose jump rates $R^{\gamma}(\eta, \xi)$ are given by

$$R^{\gamma}(\eta, \xi) = \begin{cases} R(\eta, \xi) & \text{if } \eta \text{ and } \xi \in E, \\ \gamma & \text{if } \xi = \eta^* \text{ or if } \eta = \xi^*, \\ 0 & \text{otherwise}. \end{cases}$$

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

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

$$\pi_\eta(\eta) = (1/2) \pi(\eta), \quad \pi_\eta(\eta^*) = \pi_\eta(\eta), \quad \eta \in E.$$ 

The probability measure $\pi_\eta$ is invariant for the enlarged process $\xi^{\gamma}(t)$ and is reversible whenever $\pi$ is reversible.

Let $E_N^{x,x}$, $1 \leq x \leq \kappa$, be a copy of the set $E_N^x$ and let $E_N^* = \cup_{1 \leq x \leq \kappa} E_N^{x,x}$, $E_N^{*,x} = \cup_{x \neq \xi} E_N^{x,x}$. Fix a sequence $\gamma = \gamma_N$ and denote by $\eta^\gamma(t) = \eta^{E,\gamma}$ the $\gamma$-enlargement of the trace process $\eta^{E}(t)$. Denote the generator of this Markov chain by $L_\gamma$, by $R_\gamma(\eta, \xi)$ the rate at which it jumps from $\eta$ to $\xi$, and by $\lambda_\gamma(\eta)$ the holding rates, $\lambda_\gamma(\eta) = \sum_{\xi \in E_N \cup E_N^*} R_\gamma(\eta, \xi)$. 


Denote by \( \mathbb{P}_t^\gamma \), \( \eta \in \mathcal{E}_N \cup \mathcal{E}_N^* \), the probability measure on the path space \( D(\mathbb{R}_+ \times \mathcal{E}_N \cup \mathcal{E}_N^*) \) induced by the Markov process \( \eta^*(t) \) starting from \( \eta \) and recall the definition of the hitting time and the return time introduced in (2.2). For \( x \neq y \in S \), let \( r_N(x,y) \) be the average rate at which the enlarged process \( \eta^*(t) \) jumps from \( \mathcal{E}_N^{x} \) to \( \mathcal{E}_N^{y} \):

\[
r_N(x,y) = \frac{1}{\pi_\gamma(\mathcal{E}_N^x)} \sum_{\eta \in \mathcal{E}_N^x} \pi_* (\eta) \lambda_* (\eta) \mathbb{P}_t^\gamma [H_{\mathcal{E}_N^y}^+ < H_{\mathcal{E}_N^x}^+] \\
= \frac{\gamma}{\pi_\gamma(\mathcal{E}_N^x)} \sum_{\eta \in \mathcal{E}_N^x} \pi_\gamma(\eta) \mathbb{P}_t^\gamma [H_{\mathcal{E}_N^y}^+ < H_{\mathcal{E}_N^x}^+] .
\]

(2.5)

By [1] Proposition 6.2, \( r_N(x,y) \) corresponds to the average rate at which the trace of the process \( \eta^*(t) \) on \( \mathcal{E}_N^* \) jumps from \( \mathcal{E}_N^{x} \) to \( \mathcal{E}_N^{y} \). This explains the terminology.

For two disjoint subsets \( \mathcal{A} \), \( \mathcal{B} \) of \( \mathcal{E}_N \cup \mathcal{E}_N^* \), denote by \( \text{cap}_* (\mathcal{A}, \mathcal{B}) \) the capacity between \( \mathcal{A} \) and \( \mathcal{B} \):

\[
\text{cap}_* (\mathcal{A}, \mathcal{B}) = \sum_{\eta \in \mathcal{A}} \pi_* (\eta) \lambda_* (\eta) \mathbb{P}_t^\gamma [H_{\mathcal{B}}^+ < H_{\mathcal{A}}^+] .
\]

Let \( \mathcal{A} \), \( \mathcal{B} \) be two disjoint subsets of \( S \). Taking \( \mathcal{A} = \bigcup_{x \in \mathcal{A}} \mathcal{E}_N^{x} \), \( \mathcal{B} = \bigcup_{y \in \mathcal{B}} \mathcal{E}_N^{y} \) in the previous formula, since the enlarged process may only jump from \( \eta^* \) to \( \eta \) and since \( \pi_* (\eta^*) = \pi_* (\eta) = (1/2) \pi_\gamma(\eta) \),

\[
\text{cap}_* \left( \bigcup_{x \in \mathcal{A}} \mathcal{E}_N^{x}, \bigcup_{y \in \mathcal{B}} \mathcal{E}_N^{y} \right) = \frac{\gamma}{2} \sum_{x \in \mathcal{A}} \sum_{\eta \in \mathcal{E}_N^x} \pi_\gamma(\eta) \mathbb{P}_t^\gamma [H(\bigcup_{y \in \mathcal{B}} \mathcal{E}_N^{y}) < H(\bigcup_{x \in \mathcal{A}} \mathcal{E}_N^{x})] .
\]

(2.6)

It follows from this identity and some simple algebra that

\[
\pi_\gamma(\mathcal{E}_N^x) \sum_{y \neq x} r_N(x,y) = \frac{\gamma}{2} \sum_{\eta \in \mathcal{E}_N^x} \pi_\gamma(\eta) \mathbb{P}_t^\gamma [H_{\mathcal{E}_N^x} < H_{\mathcal{E}_N^x}^+] = \text{cap}_* (\mathcal{E}_N^{x}, \mathcal{E}_N^{x}) .
\]

(2.7)

D. \( L^2 \) theory. We show in this subsection that with very few assumptions one can prove the convergence of the finite-dimensional distributions of the order \( X_t^N \). Let

\[
M_x = \min \{ \pi_\gamma(\mathcal{E}_N^x), 1 - \pi_\gamma(\mathcal{E}_N^x) \}, \quad x \in S .
\]

(2.8)

Theorem 2.2. Suppose that there exist a non-negative sequence \( \{ \theta_N : N \geq 1 \} \) and non-negative numbers \( r(x,y) \), \( x \neq y \in S \), such that

\[
\theta_N^{-1} \ll \min_{x \in S} \theta_r, \quad \lim_{N \to \infty} \theta_N r_N(x,y) = r(x,y), \quad x \neq y \in S .
\]

(L1)

Fix \( x_0 \in S \). Let \( \{ \nu_N : N \geq 1 \} \) be a sequence of probability measures concentrated on \( \mathcal{E}_N^{x_0} \), \( \nu_N(\mathcal{E}_N^{x_0}) = 1 \), and such that

\[
\mathbb{E}_\pi_\gamma \left( \left( \frac{d\nu_N}{d\pi_\gamma} \right)^2 \right) \leq \frac{C_0}{\max_{x \in S} M_x} \quad \text{(L2G)}
\]

for some finite constant \( C_0 \). Then, under \( \mathbb{P}_{\nu_N}^{\gamma} \) the finite-dimensional distributions of the time-rescaled order \( X_T^N = X_{t/\theta_N}^N \) converge to the finite-dimensional distributions of the Markov process on \( S \) which starts from \( x_0 \) and jumps from \( x \) to \( y \) at rate \( r(x,y) \).
Lemma 2.3. Assume that complement of $E$ on $E$ for all $y \in S$, $N \geq 1$, or if there is only one stable set and all the other ones have comparable measures (there exists $x \in S$ and $C_0$ such that $\lim_N \pi_\epsilon(E^y) = 1$ and $\pi_\epsilon(E^z) \leq C_0 \pi_\epsilon(E^x)$ for all $y, z \neq x$). In particular, when there are only two wells, $|S| = 2$, assumption (L2G) is satisfied by the measures $\nu_N = \pi_x$, $x = 1, 2$.

Theorem 2.4 describes the asymptotic evolution of the trace of the Markov $\eta(t)$ on $E_N$. The next lemma shows that in the time scale $\theta_N$ the time spent on the complement of $E_N$ is negligible.

**Lemma 2.3.** Assume that
\[
\lim_{N \to \infty} \frac{\pi(\Delta_N)}{\pi(E^x_N)} = 0
\]
for all $x \in S$. Let $\{\nu_N : N \geq 1\}$ be a sequence of probability measures concentrated on some well $E^x_N$, $x_0 \in S$, and satisfying (L2G). Then, for every $t > 0$,
\[
\lim_{N \to \infty} \mathbb{E}_{\nu_N} \left[ \int_0^t 1\{\eta(s\theta_N) \in \Delta_N\} ds \right] = 0. \tag{2.9}
\]

**E. Mixing theory.** If one is able to show that the process mixes inside each well before leaving the well, the assumptions on the initial state can be relaxed and the convergence of the order can be derived. Let $T_{r,x}^\text{mix}$, $x \in S$, be the mixing time of the reflected process $\eta^r(t)$.

**Theorem 2.4.** Fix $x_0 \in S$. Suppose that there exist a non-negative sequence $\{\theta_N : N \geq 1\}$ and non-negative numbers $r(x,y)$, $x \neq y \in S$, satisfying conditions (L1) and such that
\[
\lim_{N \to \infty} \sup_N \mathbb{E}_{\pi_{x_0}} [R^\epsilon(\eta, \tilde{E}^x_N)] < \infty. \tag{2.10}
\]
Let $\{\nu_N : N \geq 1\}$ be a sequence of probability measures concentrated on $E^x_N$, $\nu_N(E^x_N) = 1$. Assume that condition (L2) is fulfilled and that there exists a sequence $T_N$, $T_{r,x}^\text{mix} \ll T_N \ll \theta_N$, such that
\[
\lim_{N \to \infty} \mathbb{P}^\epsilon_{\nu_N} \left[ H_{\tilde{E}^x_N} \leq T_N \right] = 0. \tag{L4}
\]
Then, the finite-dimensional distributions of the time-rescaled order $X^N_t = X^N_{t/r_N}$ under $\mathbb{P}^\epsilon_{\nu_N}$ converges to the finite-dimensional distributions of the Markov process on $S$ which starts from $x_0$ and jumps from $x$ to $y$ at rate $r(x,y)$.

Assumption (2.10) is not difficult to be verified. By [1, Lemma 6.7],
\[
\mathbb{E}_{\pi_x} [R^\epsilon(\eta, \tilde{E}^x_N)] = \frac{1}{\pi(\tilde{E}^x_N)} \text{cap}_N(E^x_N, \tilde{E}^x_N), \tag{2.11}
\]
The Dirichlet principle [14, 15] provides a variational formula for the capacity and a bound for the expression in (2.10). We show in (3.21) below that $\sum_{y \neq x} r_N(x,y) \leq E_{\pi_x} [R^\epsilon(\eta, \tilde{E}^x_N)]$.

An assumption slightly stronger than (L4) gives tightness of the speeded-up order. For a probability measure $\nu_N$ on $E_N$, denote by $Q_{\nu_N}$ the probability measure
on the path space \( D(\mathbb{R}_+, S) \) induced by the time-rescaled order \( X_t^N = \Psi_N(\eta^N(t \theta_N)) \) starting from \( \nu_N \).

**Lemma 2.5.** Let \( \{\theta_N : N \geq 1\} \) be a sequence such that \( \theta_N^{-1} \ll \min_{x \in S} \varrho_{x,x} \) and such that for all \( x \in S \),

\[
\lim_{N \to \infty} \sup_{x} \theta_N E_{\pi_x} [R^\varrho(\eta, \tilde{\epsilon}_N^\varrho)] < \infty. \tag{2.12}
\]

Assume that there exists a sequence \( T_N \) such that \( \max_{x \in S} T_{x,x}^{\text{mix}} \ll T_N \) and such that for all \( x \in S \),

\[
\lim_{N \to \infty} \sup_{\eta \in \mathcal{E}_N^x} \mathbb{P}_\eta [H^{\tilde{\epsilon}_N^x} \leq T_N] = 0. \tag{L4U}
\]

Let \( \nu_N \) be a sequence of probability measures on \( \mathcal{E}_N \). Then, the sequence \( \{\mathbb{Q}_{\nu_N} : N \geq 1\} \) is tight.

In Section 4 we present a bound for the probability appearing in condition (L4U).

**Lemma 2.6.** Fix \( x \in S \) and suppose that there exist a set \( \mathcal{D}_N^x \subset \mathcal{E}_N^x \) and a sequence \( T_N, T_{x,x}^{\text{mix}} \ll T_N \ll \theta_N \), such that

\[
\lim_{N \to \infty} \max_{\eta \in \mathcal{D}_N^x} \mathbb{P}_\eta [H^{\tilde{\epsilon}_N^x} \leq T_N] = 0. \tag{L4E}
\]

Then, (2.9) holds for any \( t > 0 \) and any sequence of probability measures \( \nu_N \) concentrated on \( \mathcal{D}_N^x \), provided condition (L3) is in force.

Even if we are not able to prove the pointwise versions (L4U) or (L4E) of the mixing condition, we can still show that the measures of the wells converge in the Cesaro sense.

**Proposition 2.7.** Fix \( x_0 \in S \). Assume that conditions (L1), (L2) and (2.10) are fulfilled. Let \( \{\nu_N : N \geq 1\} \) be a sequence of probability measures concentrated on \( \mathcal{E}_N^x \) and satisfying conditions (2.9) and (L4). Denote by \( \{S_N(r) \mid r \geq 0\} \) the semigroup of the process \( \eta(r) \). Then, for every \( t > 0 \) and \( x \in S \),

\[
\lim_{N \to \infty} \int_0^t \left[ \nu_N S_N(\theta_N r) \right] (\tilde{\epsilon}_N) \, dr = \int_0^t \left[ \delta_{x_0} S(r) \right] (x) \, dr, \tag{2.13}
\]

where \( S(r) \) stands for the semigroup of the continuous-time Markov chain on \( S \) which jumps from \( y \) to \( z \) at rate \( r(y,z) \), and where \( \delta_{x_0} \) stands for the probability measure on \( S \) concentrated at \( x_0 \).

**F. Two valleys.** We suppose from now on that there are only two valleys, \( \mathcal{E}_N^1 = A_N \) and \( \mathcal{E}_N^2 = B_N \). It is possible in this case to establish a relation between the spectral gap of the trace process and the capacities of the enlarged process, and to re-state Theorems 2.2 and 2.4 in a simpler form. Assume that the sets \( \mathcal{E}_N^x, x = 1,2 \), have an asymptotic measure:

\[
\lim_{N \to \infty} \pi_\varrho(\tilde{\epsilon}_N^x) = m(x),
\]

and suppose, to fix ideas, that \( m(1) \leq m(2) \).
Theorem 2.8. Assume that $g \leq \min_{x=1,2} g_{r,x}$ and consider a sequence $\gamma_N$ such that $g \leq \gamma_N \leq \min_{r=1,2} g_{r,x}$. Then,

$$
\lim_{N \to \infty} \frac{\text{cap}_* \left( A_N^* \cap \mathcal{B}_N^* \right)}{g_{r,x} \pi \left( A_N \right) \pi \left( B_N \right)} = \frac{1}{2}.
$$

This result follows from [7, Theorem 2.12]. Under the assumptions of Proposition 2.1 we may replace in this statement the spectral gap of the trace process by the spectral gap of the original process. Moreover, in view of (2.7),

$$
\lim_{N \to \infty} g_{r,x}^{-1} r_N(x,y) = m(y).
$$

When there are only two valleys, the right hand side of equation (L2G) is equal to $C_0 \min \{ \pi_1(\mathcal{E}_N^1), \pi_2(\mathcal{E}_N^2) \}^{-1}$. Condition (L2G) then becomes

$$
E \left[ \left( \frac{d\nu_N}{d\pi_2} \right)^2 \right] \leq \max_{x=1,2} \frac{C_0}{\pi_2(\mathcal{E}_N^x)}
$$

for some finite constant $C_0$. The measures $\nu_N = \pi_1, \pi_2$ clearly fulfill this condition. We summarize in the next lemma the observations just made.

Lemma 2.9. Suppose that there are only two wells, $S = \{1,2\}$, and set $\theta_N = g_{r,x}^{-1}$. Then, condition (L1) is reduced to the condition that

$$
g_{r,x} \leq \min \{ g_{r,1}, g_{r,2} \}, \quad \text{(L1B)}
$$

the asymptotic rates $r(x,y)$ are given by $r(x,y) = m(y)$, and condition (L2) is always in force.

In the case of two wells, there are two different asymptotic behaviors. Assume first that $m(1) > 0$. In this case $\mathcal{E}_N^1$ and $\mathcal{E}_N^2$ are stable sets and $X_N^x$ jumps asymptotically from $x$ to $3-x$ at rate $m(3-x)$. If $m(1) = 0$ and if $\nu_N$ is a sequence of measures concentrated on $\mathcal{E}_N^1$, $\mathcal{E}_N^1$ is a metastable set, $\mathcal{E}_N^2$ a stable set, and $X_N^y$ jumps asymptotically from 1 to 2 at rate 1, remaining forever at 2 after the jump.

Remark 2.10. The average rates $r_N(x,y)$ introduced in (2.5) are different from those which appeared in [1], but can still be expressed in terms of the star-capacities:

$$
\pi_*(\mathcal{E}_N^x) r_N(x,y) = \frac{1}{2} \left\{ \text{cap}_* (\mathcal{E}_N^x, \mathcal{E}_N^y) + \text{cap}_* (\mathcal{E}_N^y, \mathcal{E}_N^x) - \text{cap}_* (\mathcal{E}_N^x \cup \mathcal{E}_N^y, \cup_{y \neq x \neq} \mathcal{E}_N^z) \right\}.
$$

To prove this identity observe that by (2.6) the right hand side is equal to

$$
\frac{\gamma}{4} \sum_{\eta \in \mathcal{E}_N^x} \pi_*(\eta) \mathbb{P}_x^\gamma [ H_{E_N^x} < H_{E_N^x}^+ ] + \frac{\gamma}{4} \sum_{\eta \in \mathcal{E}_N^y} \pi_*(\eta) \mathbb{P}_x^\gamma [ H_{E_N^x} = H_{E_N^x}^+ ].
$$

By (2.5), the first term is equal to $(1/2) \pi_*(\mathcal{E}_N^x) r_N(x,y)$. By definition of the enlarged process, the second term can be written as

$$
\frac{\gamma}{2} \sum_{\eta \in \mathcal{E}_N^x} \pi_*(\eta) \mathbb{P}_x^\gamma [ H_{E_N^x} = H_{E_N^x}^+ ].
$$

By reversibility, $\pi_*(\eta) \mathbb{P}_x^\gamma [ H_{E_N} = H_{E_N}^+ ] = \pi_*(\xi) \mathbb{P}_x^\gamma [ H_{E_N} = H_{E_N}^+ ]$, $\eta, \xi \in \mathcal{E}_N^x$. This concludes the proof of the remark.

We conclude this section pointing out an interesting difference between Markov processes exhibiting a metastable behavior and a Markov processes exhibiting the cutoff phenomena [22]. On the level of trajectories, after remaining a long time in
a metastable set, the first ones perform a sudden transition from one metastable set to another, while on the level of distributions, as stated in Proposition 2.7 below, in the relevant time scale these processes relax smoothly to the equilibrium state. In contrast, processes exhibiting the cutoff phenomena do not perform sudden transitions on the path level, but do so on the distribution level, moving quickly in a certain time scale from far to equilibrium to close to equilibrium.

3. CONVERGENCE OF THE FINITE-DIMENSIONAL DISTRIBUTIONS

We start this section with an important estimate which allows the replacement of the time integral of a function \( f : E_N \rightarrow \mathbb{R} \) by the time integral of the conditional expectation of \( f \) with respect to the \( \sigma \)-algebra generated by the partition \( E^1_N, \ldots, E^N_N \).

### 3.1. Replacement Lemma

Denote by \( \|f\|_{-1} \) the \( \mathcal{H}_{-1} \) norm associated to the generator \( L_E \) of a function \( f : E_N \rightarrow \mathbb{R} \) which has mean zero with respect to \( \pi_E \):

\[
\|f\|_{-1}^2 = \sup_h \left\{ 2 \langle f, h \rangle_{\pi_E} - \langle h, (-L_E)h \rangle_{\pi_E} \right\},
\]

where the supremum is carried over all functions \( h : E_N \rightarrow \mathbb{R} \) with finite support. By [19, Lemma 2.4], for every function \( f : E_N \rightarrow \mathbb{R} \) which has mean zero with respect to \( \pi_E \), and every \( T > 0 \),

\[
E_{\pi_E} \left[ \sup_{0 \leq t \leq T} \left( \int_0^t f(\eta^E(s)) \, ds \right)^2 \right] \leq 24 T \|f\|_{-1}^2. \tag{3.1}
\]

Similarly, for a function \( f : E^x_N \rightarrow \mathbb{R} \) which has mean zero with respect to \( \pi_x \), denote by \( \|f\|_{x,-1} \) the \( \mathcal{H}_{-1} \) norm of \( f \) with respect to the generator \( \mathcal{L}_{r,x} \) of the reflected process at \( E_N^x \):

\[
\|f\|_{x,-1}^2 = \sup_h \left\{ 2 \langle f, h \rangle_{\pi_x} - \langle h, (-\mathcal{L}_{r,x})h \rangle_{\pi_x} \right\}, \tag{3.2}
\]

where the supremum is carried over all functions \( h : E^x_N \rightarrow \mathbb{R} \) with finite support. It is clear that

\[
\sum_{x \in S} \pi_E(E^x_N) \langle h, (-\mathcal{L}_{r,x})h \rangle_{\pi_x} \leq \langle h, (-L_E)h \rangle_{\pi_E}
\]

for any function \( h : E_N \rightarrow \mathbb{R} \) with finite support. Note that the generator of the trace process \( L_E \) may have jumps from the boundary of a set \( E_N^x \) to its boundary which do not exist in the original process. There are therefore two types of contributions which appear on the right hand side but do not on the left hand side. These ones, and jumps from one set \( E_N^x \) to another. It follows from the previous inequality that for every function \( f : E_N \rightarrow \mathbb{R} \) which has mean zero with respect to each measure \( \pi_x \),

\[
\|f\|_{-1}^2 \leq \sum_{x \in S} \pi_E(E^x_N) \|f\|_{x,-1}^2. \tag{3.3}
\]

### Proposition 3.1

Let \( \{\nu_N : N \geq 1\} \) be a sequence of probability measures on \( E_N \). Then, for every function \( f : E_N \rightarrow \mathbb{R} \) which has mean zero with respect to each measure \( \pi_x \) and for every \( T > 0 \),

\[
\left( E_{\nu_N}^E \left[ \sup_{t \leq T} \left| \int_0^t f(\eta^E(s)) \, ds \right| \right] \right)^2 \leq 24 T E_{\pi_E} \left[ \left( \nu_N \pi_E \right)^2 \right] \sum_{x \in S} \pi_E(E^x_N) \|f\|_{x,-1}^2.
\]
Proof. By Schwarz inequality, the expression on the left hand side of the previous displayed equation is bounded above by
\[ E_{\pi_E} \left( \left( \frac{\nu_n}{\pi_E} \right)^2 \right) \leq E_{\pi_E} \left( \sup_{t \leq T} \left( \int_0^t f(\eta^c(s)) \, ds \right)^2 \right). \]
By (3.1) and by (3.3), the second expectation is bounded by
\[ 24 T \sum_{x \in S} \pi_E(\hat{E}_N^x) \|f\|_{x,-1}^2, \]
which concludes the proof of the proposition.

By the spectral gap, for any function \( f : \hat{E}_N^x \to \mathbb{R} \) which has mean zero with respect to \( \pi_x \), \( \|f\|_{x,-1}^2 \leq \theta_1^{-1} \langle f, f \rangle_{\pi_x} \). The next result follows from this observation and the previous proposition.

Corollary 3.2. Let \( \nu : N \geq 1 \) be a sequence of probability measures on \( E_N \). Then, for every function \( f : E_N \to \mathbb{R} \) which has mean zero with respect to each measure \( \pi_x \) and for every \( T > 0 \),
\[ \left( E_{\nu_n} \left( \sup_{t \leq T} \int_0^t f(\eta^c(s)) \, ds \right) \right)^2 \leq 24 T E_{\pi_E} \left( \left( \frac{\nu_n}{\pi_E} \right)^2 \right) \sum_{x \in S} \pi_E(\hat{E}_N^x) \theta_1^{-1} \langle f, f \rangle_{\pi_x}. \]

We have seen in (2.7) that \( \text{cap}_*(\hat{E}_N^x, \hat{E}_N^x) = \pi_x(\hat{E}_N^x) \sum_{y \neq x} r_N(x, y) \). For similar reasons, \( \text{cap}_*(\hat{E}_N^x, \hat{E}_N^x) = \pi_x(\hat{E}_N^x) r_N(x, y) \). If \( \theta_1 r_N(x, y) \) converges, as postulated in assumption (L1), we obtain from these identities that
\[ \text{cap}_*(\hat{E}_N^x, \hat{E}_N^x) \leq C_0 \theta_1^{-1} M_x \tag{3.4} \]
for some finite constant \( C_0 \), where \( M_x \) has been introduced in (2.8).

The equilibrium potentials. Fix a sequence \( \gamma = \gamma_N \) such that \( \theta_1^{-1} \ll \gamma \ll \min_{x \in S} \theta_1^{-1} \theta_{1,x} \) and recall that we denote by \( \eta^c(t) \) the \( \gamma \)-enlargement of the trace process \( \eta^c(t) \). Denote by \( V_x, x \in S \), the equilibrium potential between the sets \( \hat{E}_N^x \) and \( \hat{E}_N^x \), \( V_x(\eta) = \mathbb{P}^{\eta, \gamma}_x[H_{\hat{E}_N^x} < H_{\hat{E}_N^x}] \). Since \( L_x V_x = 0 \) on \( E_N \), we deduce that
\[ (L_x V_x)(\eta) = -\gamma [1 - V_x(\eta)], \quad \eta \in E_N^x, \]
\[ (L_x V_x)(\eta) = \gamma V_x(\eta), \quad \eta \in \hat{E}_N^x. \tag{3.5} \]

Moreover, since \( \pi_x(\eta) = (1/2) \pi_E(\eta), \eta \in E_N^x \),
\[ \text{cap}_* (\hat{E}_N^x, \hat{E}_N^x) = \]
\[ \frac{1}{2} \left\{ \gamma \sum_{\eta \in E_N^x} \pi_E(\eta) [1 - V_x(\eta)]^2 + \langle (-L_x)V_x, V_x \rangle_{\pi_E} + \gamma \sum_{\eta \in \hat{E}_N^x} \pi_E(\eta) V_x(\eta)^2 \right\}. \tag{3.6} \]

By assumption (L1), for all \( x \neq y \in S \), \( r_N(x, y) \leq C_0 \theta_1^{-1} \) for some finite constant \( C_0 \) and for all \( N \) large enough. Hence, by (3.4) and by (3.6), for all \( x \in S \)
\[ \gamma \sum_{\eta \in E_N^x} \pi_E(\eta) [1 - V_x(\eta)]^2 + \langle (-L_x)V_x, V_x \rangle_{\pi_E} + \gamma \sum_{\eta \in \hat{E}_N^x} \pi_E(\eta) V_x(\eta)^2 \leq \frac{C_0 M_x}{\theta_1}. \tag{3.7} \]
Uniqueness of limit points. Recall the definition of the measure $Q_{\nu_N}$ introduced just before Lemma 2.5 and let $\mathfrak{L}$ be the generator of the $S$-valued Markov process given by

$$(\mathfrak{L} F)(x) = \sum_{y \in S} r(x,y)[F(y) - F(x)].$$

**Proposition 3.3.** Assume that the hypotheses of Theorem 2.2 are in force. Then, the sequence $Q_{\nu_N}$ has at most one limit point, the probability measure on $D(\mathbb{R}_+, S)$ induced by the Markov process with generator $\mathfrak{L}$ starting from $x_0$.

**Proof.** To prove the uniqueness of limit points, we use the martingale characterization of Markov processes. Fix a function $F : S \to \mathbb{R}$ and a limit point $Q_*$ of the sequence $Q_{\nu_N}$. We claim that

$$M_t^F := F(X_t) - F(X_0) - \int_0^t (\mathfrak{L} F)(X_s) \, ds$$

is a martingale under $Q_*$. Fix $0 \leq s < t$ and a bounded function $U : D(\mathbb{R}_+, S) \to \mathbb{R}$ depending only on $\{X_r : 0 \leq r \leq s\}$ and continuous for the Skorohod topology. We shall prove that

$$\mathbb{E}_{Q_*} [M_t^F U] = \mathbb{E}_{Q_*} [M_s^F U].$$

Let $G(\eta) = \sum_{x \in S} F(x)V_x(\eta)$, $\eta \in \mathcal{E}_N$. By the Markov property of the trace process $\eta^E(t)$,

$$M_t^N := G(\eta^E(t\theta_N)) - G(\eta^E(0)) - \int_0^{t\theta_N} (L\mathcal{E}G)(\eta^E(s)) \, ds$$

is a martingale. Let $U^N := U(X_N)$. As $\{M_t^N : t \geq 0\}$ is a martingale,

$$\mathbb{E}_{\nu_N}^E [M_t^N U^N] = \mathbb{E}_{\nu_N}^E [M_s^N U^N]$$

so that

$$\mathbb{E}_{\nu_N}^E \left[ U^N \left\{ G(\eta^E(t\theta_N)) - G(\eta^E(s\theta_N)) - \int_{s\theta_N}^{t\theta_N} (L\mathcal{E}G)(\eta^E(r)) \, dr \right\} \right] = 0.$$ (3.10)

**Claim A:** For all $x \in S$,

$$\lim_{N \to \infty} \sup_{t \geq 0} \mathbb{E}_{\nu_N}^E \left[ |1_{E_N^x}(\eta^E(t\theta_N)) - V_x(\eta^E(t\theta_N))| \right] = 0.$$

Indeed, denote by $S_{\xi}(t)$, $t \geq 0$, the semigroup associated to the trace process $\eta^E(t)$, and by $h_t$ the Radon-Nikodym derivative $d\nu_N S_{\xi}(t)/d\pi_{\xi}$. It is well known that $E_{\pi_{\xi}} [h_t^2] \leq E_{\pi_{\xi}} [h_0^2]$. Hence, by Schwarz inequality, the square of the expectation appearing in the previous displayed formula is bounded above by

$$E_{\pi_{\xi}} \left[ \left( \frac{d\nu_N}{d\pi_{\xi}} \right)^2 \right] E_{\pi_{\xi}} \left[ |1_{E_N^x} - V_x|^2 \right].$$

To conclude the proof of the claim it remains to recall the definition of the sequence $\nu^*$, the estimate (3.7) and the assumption on the sequence of probability measures $\nu_N$.

It follows from Claim A that

$$\lim_{N \to \infty} \sup_{t \geq 0} \mathbb{E}_{\nu_N}^E \left[ |(F \circ \Psi)(\eta^E(t\theta_N)) - G(\eta^E(t\theta_N))| \right] = 0.$$
Therefore, by (3.10),

$$\lim_{N \to \infty} E^\nu_{\nu_N} \left[ U^N \left\{ \Delta_{s,t} F - \int_{\theta_N}^{t \theta_N} (L_{x,\nu}) (\nu (r)) \, dr \right\} \right] = 0.$$  \hspace{0.5cm} (3.11)

where $\Delta_{s,t} F = (F \circ \nu (s \theta_N)) - (F \circ \nu (t \theta_N)) = F(X^N_{s \theta_N}) - F(X^N_{t \theta_N})$.

**Claim B:** Denote by $P$ the $\sigma$-algebra generated by the partition $\mathcal{E}^N_N$, $z \in S$. For all $T > 0$, $x \in S$,

$$\lim_{N \to \infty} E^\nu_{\nu_N} \left[ \sup_{t \leq T} \left| \int_0^t \left\{ (L_{x,\nu}) (\nu (s)) - E[L_{x,\nu} | \mathcal{P}] (\nu (s)) \right\} \, ds \right| \right] = 0.$$  \hspace{0.5cm} (3.12)

for some finite constant $C_0$, where $C$ stands for $G - E_{\nu_N} [G]$. By (3.5), on the set $\mathcal{E}^N_x, L_{x,\nu} = -\gamma [1 - V_x (\nu)]$. Hence, by the spectral gap on the reflected process and by (3.7),

$$\| L_{x,\nu} V_x \|_{\nu_N}^2 = \gamma^2 \| 1 - V_x \|_{\nu_N}^2 \leq \frac{\gamma^2}{\theta_{r,x}} \| 1 - V_x \|_{\nu}^2 \leq \frac{C_0 \gamma M_x}{\pi \nu (E^N_N) \theta_{r,x} \theta_N}$$

for some finite constant $C_0$. Similarly, since $L_{x,\nu} V_x = \gamma V_x (\nu)$ on the set $\mathcal{E}^N_y$, $y \neq x$,

$$\| L_{x,\nu} V_x \|_{\nu_N}^2 \leq \frac{C_0 \gamma M_x}{\pi \nu (E^N_N) \theta_{r,y} \theta_N}.$$  \hspace{0.5cm} (3.13)

Therefore, the sum appearing in (3.12) is bounded by $C_0 T |S| \gamma \max_{x \in S} \theta_{r,x}^{-1}$ which vanishes as $N \uparrow \infty$ by definition of $\gamma$, proving Claim B.

It follows from (3.11) and Claim B that

$$\lim_{N \to \infty} E^\nu_{\nu_N} \left[ U^N \left\{ \Delta_{s,t} F - \int_{\theta_N}^{t \theta_N} E[L_{x,\nu} | \mathcal{P}] (\nu (r \theta_N)) \, dr \right\} \right] = 0.$$  \hspace{0.5cm} (3.14)

We affirm that

$$E[L_{x,\nu} | \mathcal{P}] (\nu) = \sum_{x \in S} 1_{\{ \nu \in \mathcal{E}^N_x \}} \sum_{y \in S} r_N (x, y) [F (y) - F (x)].$$  \hspace{0.5cm} (3.15)

Indeed, by (3.5),

$$E[L_{x,\nu} V_x | \mathcal{P}] = \begin{cases} -\gamma \sum_{y \in S} \frac{\pi \nu (\eta)}{\pi \nu (E^N_N)} \mathbb{P}^\eta \nu (H_{E^N_N} \leq H_{E^N_N}) \nu, & \eta \in \mathcal{E}^N_x, \\ \gamma \sum_{y \in S} \frac{\pi \nu (\eta)}{\pi \nu (E^N_N)} \mathbb{P}^\eta \nu (H_{E^N_N} > H_{E^N_N}) \nu, & \eta \in \mathcal{E}^N_y, y \neq x. \end{cases}$$

By (2.7), on the set $\mathcal{E}^N_x, E[L_{x,\nu} V_x | \mathcal{P}] = - \sum_{y \neq x} r_N (x, y)$, and by (2.5), on the set $\mathcal{E}^N_y, E[L_{x,\nu} V_x | \mathcal{P}] = r_N (y, x)$. To conclude the proof of (3.14) it remains to recall the definition of $G$.

By (3.13), (3.14) and by definition of $X^N_t$,

$$\lim_{N \to \infty} E^\nu_{\nu_N} \left[ U^N \left\{ \Delta_{s,t} F - \int_{\theta_N}^{t \theta_N} \sum_{y \in S} r_N (X^N_{r \theta_N}, y) [F (y) - F (X^N_{r \theta_N})] \, dr \right\} \right] = 0.$$
Since $\Delta_{t\theta} F = F(X_{t\theta}^N) - F(X_{s\theta}^N)$, since $U$ has been assumed to be continuous for the Skorohod topology and since $Q_s$ is a limit point of the sequence $Q_{t\theta}$, by assumption (L1)

$$E_{Q_s} \left[ U \left( F(X_t) - F(X_s) - \int_s^t \sum_{y \in S} r(X_r,y) [F(y) - F(X_r)] \, dr \right) \right] = 0,$$

proving (3.8) and the proposition.

**Proof of Theorem 2.2.** The proof is similar to the one of Proposition 3.3. We prove the convergence of the one-dimensional distributions. The extension to higher dimensional distributions is clear. Fix a function $F : S \to \mathbb{R}$. We claim that for every $T \geq 0$,

$$\limsup_{N \to \infty} \sup_{0 \leq t \leq T} \left| \mathbb{E}_{\nu_N} \left[ F(X_{t\theta}^N) - F(X_{s\theta}^N) - \int_s^t (\mathcal{L}F)(X_{r\theta}^N) \, dr \right] \right| = 0.$$  \hspace{1cm} (3.15)

Recall the definition of the function $G : \mathbb{E}_N \to \mathbb{R}$ introduced in the proof of the previous proposition. By Claim A and since $G(\eta^f(t\theta_N)) - \int_0^t \theta_N(LfG)(\eta^f(s\theta_N)) \, ds$ is a martingale, to prove (3.15), it is enough to show that

$$\limsup_{N \to \infty} \sup_{0 \leq t \leq T} \left| \mathbb{E}_{\nu_N} \left[ \int_0^t (\mathcal{L}F)(X_{r\theta}^N) \, dr - \int_0^t (\mathcal{L}F)(X_{r\theta}^N) \, dr \right] \right| = 0.$$  \hspace{1cm} (3.15)

By Claim B, by the identity (3.11) and by the definition of $X^N_t$, the proof of (3.15) is further reduced to the proof that

$$\limsup_{N \to \infty} \sup_{0 \leq t \leq T} \left| \mathbb{E}_{\nu_N} \left[ \int_0^t (\mathcal{L}F)(X_{r\theta}^N) \, dr - \int_0^t (\mathcal{L}F)(X_{r\theta}^N) \, dr \right] \right| = 0,$$

where $(\mathcal{L}F)(x) = \sum_{y \in S} \theta_N r_N(x,y) [F(y) - F(x)]$. To conclude the proof of (3.15), it remains to recall assumption (L1).

It follows from (3.15) that the sequence $f_N(t) = \mathbb{E}_{\nu_N} [F(X_{t\theta}^N)]$ is equicontinuous in any compact interval $[0, T]$. Moreover, if $F$ is an eigenfunction of the operator $\mathcal{L}$ associated to an eigenvalue $\lambda$, all limit points $f(t)$ of the subsequence $f_N(t)$ are such that

$$f(t) - F(x_0) = \int_0^t \lambda f(r) \, dr, \quad 0 \leq t \leq T,$$

which yields uniqueness of limit points.

**Proof of Theorem 2.4.** Recall that $T^\text{mix}_{r,s} : x \in S$, stands for the mixing time of the reflected process $\eta^f x(t)$. We prove that the one-dimensional distributions converge. The extension to higher dimensional distributions is straightforward. In view of Theorem 2.2, it is enough to show that for each function $F : S \to \mathbb{R}$,

$$\lim_{N \to \infty} \left| \mathbb{E}_{\nu_N}^F \left[ F(X_{t\theta}^N) \right] - \mathbb{E}_{\pi_0}^F \left[ F(X_{t\theta}^N) \right] \right| = 0.$$  \hspace{1cm} (3.16)

Let $T_N$ be a sequence satisfying the assumptions of the theorem. We may write $\mathbb{E}_{\nu_N}^F \left[ F(X_{t\theta}^N) \right] as

$$\mathbb{E}_{\nu_N}^F \left\{ 1 \{ H_{\nu_N}^\theta > T_N \} F(X_{t\theta}^N) \right\} + \mathbb{E}_{\nu_N}^\theta \left\{ 1 \{ H_{\nu_N}^\theta \leq T_N \} F(X_{t\theta}^N) \right\}.$$

The second term is absolutely bounded by $C_0 \mathbb{E}_{\nu_N}^\theta \left\{ H_{\nu_N}^\theta \leq T_N \right\}$ for some finite constant $C_0$ independent of $N$ and which may change from line to line. By hypothesis,
this latter probability vanishes as $N \uparrow \infty$. By the Markov property, the first term in the previous displayed equation is equal to
\[
\mathbb{E}_{\nu_N}^E \left[ 1\{H_{\xi_N} > T_N\} \mathbb{E}_{\eta(T_N)}^E \left[ F(X_{t\theta_N-T_N}^N) \right] \right].
\]
On the set $\{H_{\xi_N} > T_N\}$ we may couple the trace process with the reflected process in such a way that $\eta^E(t) = \eta^{x_0}(t)$ for $t \leq T_N$. The previous expectation is thus equal to
\[
\mathbb{E}_{\nu_N}^E \left[ \mathbb{E}_{\eta_T^{x_0}(T_N)}^E \left[ F(X_{t\theta_N-T_N}^N) \right] \right] - \mathbb{E}_{\nu_N}^E \left[ 1\{H_{\xi_N}^{x_0} \leq T_N\} \mathbb{E}_{\eta_T^{x_0}(T_N)}^E \left[ F(X_{t\theta_N-T_N}^N) \right] \right].
\]
As before, the second term vanishes as $N \uparrow \infty$. The first expectation is equal to
\[
\mathbb{E}_{\pi_{x_0}}^E \left[ F(X_{t\theta_N-T_N}^N) \right] + R_N(t),
\]
where $R_N(t)$ is absolutely bounded by $C_0\|\nu_N S^{x_0}(T_N) - \pi_{x_0}\|_{TV}$. In this formula, $\|\mu - \nu\|_{TV}$ stands for the total variation distance between $\mu$ and $S^{x}(t)$ represents the semi-group of the reflected process. By definition of the mixing time, this last expression is less than or equal to $(1/2)^{T_N/r^{mix}}$, which vanishes as $N \uparrow \infty$ by assumption.

At this point we repeat the same argument with the measure $\nu_N$ replaced by the local equilibrium $\pi_{x_0}$. To estimate $\mathbb{P}^E_{\pi_{x_0}}[H_{\xi_N}^{x_0} \leq T_N]$ we write this expression as
\[
\sum_{\eta \in \mathcal{E}_{\xi_N}^{x_0}} \left\{ \pi_{x_0}(\eta) - \pi_{x_0}^*(\eta) \right\} \mathbb{P}_{\pi_{x_0}}^E[H_{\xi_N}^{x_0} \leq T_N] + \mathbb{P}_{\pi_{x_0}}^E[H_{\xi_N}^{x_0} \leq T_N],
\]
where $\pi_{x_0}^*$ is the quasi-stationary measure associated to the trace process $\eta^E(t)$ killed when it hits $\xi_N^{x_0}$. The first term is less than or equal to
\[
\sum_{\eta \in \mathcal{E}_{\xi_N}^{x_0}} \left| \frac{\pi_{x_0}(\eta)}{\pi_{x_0}^*(\eta)} - 1 \right| \leq \left( \sum_{\eta \in \mathcal{E}_{\xi_N}^{x_0}} \pi_{x_0}(\eta) \left( \frac{\pi_{x_0}^*(\eta)}{\pi_{x_0}(\eta)} - 1 \right)^2 \right)^{1/2}.
\]
By Proposition 2.1, (17) and Lemma 2.2 in [7], the expression inside the square root on the right hand side of the previous formula is bounded by $\varepsilon_{x_0}/[1 - \varepsilon_{x_0}]$, where $\varepsilon_{x_0} = E_{\pi_{x_0}}[R^E(\eta, \xi_N^{x_0})]/\|\nu_{x_0}\|$. By (2.10), $\varepsilon_{x_0} \leq C_0(\theta_N \varepsilon_{x_0})^{-1}$ for some finite constant $C_0$ and by hypothesis, $\theta_N^{-1} \ll \varepsilon_{x_0}$. This shows that the first term in (3.17) vanishes as $N \uparrow \infty$.

Finally, since $\pi_{x_0}^*$ is the quasi-stationary state, under $\mathbb{P}_{\pi_{x_0}^*}$, the hitting time of $\xi_N^{x_0}$, $H_{\xi_N}$, has an exponential distribution whose parameter we denote by $\phi_{x_0}^*$. By (7) Lemma 2.2, $\phi_{x_0}^*$ is bounded by $E_{\pi_{x_0}}[R^E(\eta, \xi_N^{x_0})] \leq C_0/\theta_N$, for some finite constant $C_0$.

Hence,
\[
\mathbb{P}_{\pi_{x_0}^*}^E[H_{\xi_N} \leq T_N] = 1 - e^{-\phi_{x_0}^* T_N} \leq 1 - e^{-C_0(T_N/\theta_N)},
\]
an expression which vanishes as $N \uparrow \infty$. \hfill \Box

**Proof of Lemma 2.3.** Let $\nu_N$ be a sequence of probability measures satisfying (L2G). By Schwarz inequality, the square of the expectation appearing in the statement of the lemma is bounded above by
\[
\frac{1}{\pi(E_N)} E_{\pi_{x_0}^*} \left[ \left( \frac{d\nu_N}{d\pi_{x_0}^*} \right)^2 \right] \mathbb{E}_{\pi_{x_0}^*} \left[ \left( \int_0^t 1\{\eta(s) \in \Delta_N\} ds \right)^2 \right].
\]
Markov property, the expectation appearing in (2.9) with $\nu_{L^2G}$.

By assumption (Schwarz inequality, the second expectation is less than or equal to $\eta_{\pi_N}$.) On the other hand, by obvious modifications. Consider a sequence of initial states $\eta$ with $\nu = \delta_{\eta_N}$ is bounded above by

$$t \mathbb{E}_\pi \left[ \int_0^t 1\{\eta(s\theta_N) \in \Delta_N\} ds \right] = t^2 \pi(\Delta_N),$$

which concludes the proof. □

**Proof of Lemma 2.6.** The proof of this result is similar to the previous one with Schwarz inequality. Consider a sequence of initial states $\eta_N$ in $\mathcal{D}_N$. By the Markov property, the expectation appearing in (2.9) with $\nu_N = \delta_{\eta_N}$ is bounded above by

$$\mathbb{E}_{\eta_N} \left[ 1\{H(T_N^\pi) > T_N\} \mathbb{E}_{\eta(T_N)} \left[ \int_0^t 1\{\eta(s\theta_N) \in \Delta_N\} ds \right] \right] + T_N/\theta_N + t \mathbb{P}_{\eta_N} \left[ H(T_N^\pi) \leq T_N \right],$$

where we replaced $t - T_N$ by $t$ in the time integral. By assumption, the second and the third term vanish as $N \uparrow \infty$. On the set $\{H(T_N^\pi) > T_N\}$ we may replace $\eta(T_N)$ by $\eta^{r,T_N^\pi}(T_N)$, where $\eta^{r,T_N^\pi}(t)$ stands for the process $\eta(t)$ reflected at $T_N^\pi$. After this replacement, we may remove the indicator and estimate the expectation by

$$t \left\| \delta_N \right\|_{ST,T_N^\pi} - \pi_{TV} + \mathbb{E}_{\pi_{TV}} \left[ \int_0^t 1\{\eta(s\theta_N) \in \Delta_N\} ds \right],$$

where $ST,T_N^\pi$ represents the semigroup of the reflected process $\eta^{r,T_N^\pi}(t)$ and $\pi_{TV}$ is bounded by $\pi(T_N)/\pi(\mathcal{F}_N^\pi)$, which vanishes in view of condition (L3).

**Proof of Proposition 2.7.** The proof of this proposition relies on a comparison between the original process and the trace process presented below in equations (3.15) and (3.19). Let $\{T_\epsilon(t) \mid t \geq 0\}$ be the time spent on the set $\mathcal{E}_N$ by the process $\eta(s)$ in the time interval $[0,t]$,

$$T_\epsilon(t) = \int_0^t 1\{\eta(s) \in \mathcal{E}_N\} ds.$$

Denote by $S_\epsilon(t)$ the generalized inverse of $T_\epsilon(t)$. $S_\epsilon(t) = \sup\{s \geq 0 \mid T_\epsilon(s) \leq t\}$, and recall that the trace process is defined as $\eta^\pi(t) = \eta(S_\epsilon(t))$.

By definition of the trace process, for every $t \geq 0$,

$$\int_0^t 1\{\eta(s) \in \mathcal{E}_N\} ds \leq \int_0^t 1\{\eta^\pi(s) \in \mathcal{E}_N\} ds. \quad (3.18)$$

On the other hand,

$$\int_0^t 1\{\eta^\pi(r) \in \mathcal{E}_N\} dr = \int_0^t 1\{\eta(S_\epsilon(r)) \in \mathcal{E}_N\} dr.$$

By a change of variables, the previous integral is equal to

$$\int_0^{S_\epsilon(t)} 1\{\eta(r) \in \mathcal{E}_N\} dr.$$

Let $T_\Delta(t)$, $t \geq 0$, be the time spent by the process $\eta(s)$ on the set $\Delta_N$ in the time interval $[0,t]$, $T_\Delta(t) = \int_0^t 1\{\eta(s) \in \Delta_N\} ds$. Since $T_\epsilon(t) + T_\Delta(t) = t$, on the set
$T_{\Delta}(t_{0}) < \delta$ and for $t \leq t_{0} - \delta$, $T_{\Delta}(t + \delta) > t$, so that $S_{\xi}(t) \leq t + \delta$. Putting together all previous estimates we get that on the set $T_{\Delta}(t_{0}) < \delta$ and for $t \leq t_{0} - \delta$,

$$
\int_{0}^{t} 1\{\eta^c(s) \in \mathcal{E}_N^c\} \, ds \leq \int_{0}^{t+\delta} 1\{\eta(s) \in \mathcal{E}_N^x\} \, ds . \quad (3.19)
$$

We turn now to the proof of the proposition. We may rewrite the time integral appearing on the left hand side of (2.13) as

$$
E_{\nu_N}\left[ \int_{0}^{t} 1\{\eta(r\theta_N) \in \mathcal{E}_N^x\} \, dr \right] . \quad (3.20)
$$

By (3.18), this expectation is bounded above by

$$
E_{\nu_N}\left[ \int_{0}^{t} 1\{\eta^c(r\theta_N) \in \mathcal{E}_N^c\} \, dr \right] = E_{\nu_N}\left[ \int_{0}^{t} 1\{X_N^{r\theta_N} = x\} \, dr \right] .
$$

By Theorem 2.4, the right hand side converges as $N \uparrow \infty$ to the right hand side of (2.13).

Fix $\delta > 0$. The expectation (3.20) is bounded below by

$$
E_{\nu_N}\left[ \{T_{\Delta}(r\theta_N) < \delta\theta_N\} \int_{0}^{t} 1\{\eta(r\theta_N) \in \mathcal{E}_N^x\} \, dr \right].
$$

By (3.19), this expression is bounded below by

$$
E_{\nu_N}\left[ \{T_{\Delta}(r\theta_N) < \delta\theta_N\} \int_{0}^{t-\delta} 1\{\eta^c(r\theta_N) \in \mathcal{E}_N^c\} \, dr \right]
\geq E_{\nu_N}\left[ \int_{0}^{t-\delta} 1\{\eta^c(r\theta_N) \in \mathcal{E}_N^c\} \, dr \right] - t E_{\nu_N}[T_{\Delta}(r\theta_N) \geq \delta\theta_N] .
$$

By (2.41), the second term vanishes as $N \uparrow \infty$, while by Theorem 2.4 the second one converges to the right hand side of (2.13) as $N \uparrow \infty$ and then $\delta \downarrow 0$.

\[\Box\]

\textbf{The jump rates.} Recall the definition (2.3) of the rates $r_N(x, y)$. For all $x \in S$,

$$
\sum_{y \neq x} r_{N}(x, y) \leq E_{\pi_x}[R^E(\eta, \tilde{\mathcal{E}}_N)] . \quad (3.21)
$$

Indeed, by (2.7) and by the Dirichlet principle,

$$
\pi_{x}(\mathcal{E}_N^x) \sum_{y \neq x} r_{N}(x, y) = \text{cap}_{x}(\mathcal{E}_N^{x, x}, \tilde{\mathcal{E}}_N^{x, x}) = \inf_{f}((-L_{*}f), f)_{\pi_x} ,
$$

where the infimum is carried over all functions $f : \mathcal{E}_N \cup \mathcal{E}_N^{x} \rightarrow \mathbb{R}$ equal to 1 on $\mathcal{E}_N^{x, x}$ and equal to 0 on $\tilde{\mathcal{E}}_N^{x, x}$. Taking $f = 1\{\mathcal{E}_N^{x} \cup \mathcal{E}_N^{x, x}\}$ and computing the Dirichlet form of this function we get (3.21).

\section{4. On Assumptions (L4) and (L4U)}

We present in this section two estimates of $\mathbb{P}_{\eta}^{\xi}[H_{\tilde{\mathcal{E}}_N^{x}} \leq T_N]$. We start with a bound of this probability in terms of an equilibrium potential. Denote by $W_{x, \gamma}$, $x \in S$, $\gamma > 0$, the equilibrium potential between $\tilde{\mathcal{E}}_N^{x} \cup \tilde{\mathcal{E}}_N^{x, x}$ and $\mathcal{E}_N^{x}$ for the $\gamma$-enlargement of the trace process $\eta(t)$:

$$
W_{x, \gamma}(\eta) = \mathbb{P}_{\eta}^{\xi, \gamma}[H_{\tilde{\mathcal{E}}_N^{x}} \cup \tilde{\mathcal{E}}_N^{x, x} < H_{\tilde{\mathcal{E}}_N^{x}} < H_{\mathcal{E}_N^{x}}] = \mathbb{P}_{\eta}^{\xi, \gamma}[H_{\tilde{\mathcal{E}}_N^{x}} < H_{\mathcal{E}_N^{x}}] .
$$
Lemma 4.1. Fix $x \in S$. Then, for all $\eta \in \mathcal{E}_N^x$, $\gamma > 0$ and $A > 0$,
\[
P^\mathcal{E}_N[H_{\tilde{\mathcal{E}}_N^x} \leq \gamma^{-1}] \leq e W^*_{x,\gamma}(\eta),
\]
\[
W^*_{x,\gamma}(\eta) - \frac{e^{-A}}{1 - e^{-A}} \leq P^\mathcal{E}_N[H_{\tilde{\mathcal{E}}_N^x} \leq A\gamma^{-1}].
\]

Proof. Fix $x \in S$. By definition of the equilibrium potential,
\[
\begin{cases}
(L_* W^*_{x,\gamma})(\eta) = 0 & \eta \in \mathcal{E}_N^x, \\
W^*_{x,\gamma}(\eta) = 1 & \eta \in \mathcal{E}_N^x \cup \mathcal{E}_N^{x*}, \\
W^*_{x,\gamma}(\eta) = 0 & \eta \in \mathcal{E}_N^{x*}.
\end{cases}
\]
By definition of the generator $L_*$ and since the equilibrium potential $W^*_{x,\gamma}$ vanishes on the set $E_N^{x*}$, on the set $\mathcal{E}_N^x$, we have that
\[
(L_* W^*_{x,\gamma})(\eta) = \gamma W^*_{x,\gamma}(\eta), \quad \eta \in \mathcal{E}_N^x.
\]
Since $W^*_{x,\gamma}$ is equal to 1 on the set $\mathcal{E}_N^x$, we conclude that
\[
W^*_{x,\gamma}(\eta) = E^\mathcal{E}_N[\exp\{-\gamma H_{\tilde{\mathcal{E}}_N^x}\}], \quad \eta \in \mathcal{E}_N.
\]
On the other hand, by Tchebychev inequality and by the previous identity,
\[
P^\mathcal{E}_N[H_{\tilde{\mathcal{E}}_N^x} \leq \gamma^{-1}] = E^\mathcal{E}_N[e^{-\gamma H_{\tilde{\mathcal{E}}_N^x}} \geq e^{-1}] \leq e E^\mathcal{E}_N[e^{-\gamma H_{\tilde{\mathcal{E}}_N^x}}] = e W^*_{x,\gamma}(\eta).
\]
Conversely, fix $A > 0$ and let $T(\gamma)$ be an exponential time of parameter $\gamma$ independent of the trace process $\eta^\mathcal{E}(s)$. It is clear that for $\eta \in \mathcal{E}_N^x$,
\[
W^*_{x,\gamma}(\eta) = P^\mathcal{E}_N[H_{\mathcal{E}_N^x} < H_{\mathcal{E}_N^{x*}}] = P^\mathcal{E}_N[H_{\mathcal{E}_N^x} < T(\gamma)].
\]
By definition of $T(\gamma)$, the last probability is equal to
\[
\int_0^\infty P^\mathcal{E}_N[H_{\mathcal{E}_N^x} < t] \gamma e^{-\gamma t} dt \leq \frac{e^{-A}}{1 - e^{-A}}.
\]
An elementary computation permits to conclude the proof of the lemma.

The second assertion of the previous lemma shows that we do not lose much in the first one.

Corollary 4.2. Let $\nu_N$ be a probability measure concentrated on the set $\mathcal{E}_N^x$. Then, for all $\gamma > 0$,
\[
E^\mathcal{E}_N[\gamma^{-1}]^2 \leq \frac{2e^2}{\gamma} E^\mathcal{E}_x \left[ \left( \frac{\nu_N}{\pi^x} \right)^2 \right] \sum_{x,\gamma} \nu_N(\eta) \rho^x(\gamma) \cap_{x} (\mathcal{E}_N^x, \mathcal{E}_N^{x*}).
\]

Proof. Recall that we denote by $\eta^*$ the copy of the state $\eta$. By definition of the enlarged process and by Schwarz inequality,
\[
P^\mathcal{E}_N[H_{\mathcal{E}_N^x} < H_{\mathcal{E}_N^{x*}}] = \sum_{\eta \in \mathcal{E}_N^x} \nu_N(\eta) P^\mathcal{E}_N[H_{\mathcal{E}_N^x} < H_{\mathcal{E}_N^{x*}}]
\]
\[
\leq \left\{ E^\mathcal{E}_x \left[ \left( \frac{\nu_N}{\pi^x} \right)^2 \right] \sum_{\eta \in \mathcal{E}_N^x} \pi^x(\eta) P^\mathcal{E}_N[H_{\mathcal{E}_N^x} < H_{\mathcal{E}_N^{x*}}] \right\}^{1/2}.
\]
In the previous sum we may replace $\pi^x(\eta)$ by $2\pi^* (\eta^*)$. After the replacement, the sum becomes $2(1 - \gamma^{-1}) \cap_{x} (\mathcal{E}_N^x, \mathcal{E}_N^{x*})$. This estimate together with Lemma 4.1 concludes the proof of the corollary.
Comments on assumption (L4U). We present in this subsection two strategies to prove that the equilibrium potential $W_{x,\gamma}^*(\eta)$ vanishes. We apply the first technique in Example A of Section 8.

A. Monotonicity. On the one hand, it is always possible to couple two trace processes $\eta^\varepsilon(t)$ starting from different initial states in such a way that both reach the set $\tilde{E}^*_N$ at the same time. Assume that the equilibrium potential $W_{x,\gamma}^*$ satisfy some property $\mathcal{P}$. For example, if the state space $E_N$ is partially ordered and if the process $\eta^\varepsilon(t)$ is monotone, the equilibrium potential might be monotone. By the Dirichlet principle,

$$\text{cap}_{\pi}(\tilde{E}^{*,x}_N, \tilde{E}^{*,x}_N + \tilde{E}^{*,x}_N) = \langle W_{x,\gamma}^*(\eta), (-L_*)W_{x,\gamma}^*(\eta) \rangle_{\pi_*} = \inf_{f} \langle f, (-L_*)f \rangle_{\pi_*},$$

where the supremum is carried over all functions $f$ vanishing at $\tilde{E}^{*,x}_N$, equal to 1 on $\tilde{E}^{*,x}_N$ and satisfying condition $\mathcal{P}$. Fix a configuration $\eta \in \tilde{E}^{*,x}_N$ and denote by $R_N(\varepsilon), \varepsilon > 0$, the right hand side of the previous formula when we impose the further restriction that $f(\eta) \geq \varepsilon$.

To prove that $W_{x,\gamma}^*(\eta)$ vanishes as $N \uparrow \infty$, it is enough to show that for every $\varepsilon > 0$, $\text{cap}_{\pi}(\tilde{E}^{*,x}_N, \tilde{E}^{*,x}_N + \tilde{E}^{*,x}_N) \ll R_N(\varepsilon)$. Indeed, suppose by contradiction that $W_{x,\gamma}^*(\eta)$ does not vanish as $N \uparrow \infty$. There exists in this case $\varepsilon > 0$ and a subsequence $N_j$, still denoted by $N$, for which $W_{x,\gamma}^*(\eta) \geq \varepsilon$ for all $N$. Therefore,

$$R_N(\varepsilon) \leq \langle W_{x,\gamma}^*, (-L_*)W_{x,\gamma}^* \rangle_{\pi_*} = \text{cap}_{\pi}(\tilde{E}^{*,x}_N, \tilde{E}^{*,x}_N + \tilde{E}^{*,x}_N),$$

proving our claim.

B. Capacities. To present the second form of estimating the equilibrium potential, we start with a general result which expresses the equilibrium potential as a ratio between capacities. Consider a reversible Markov chain $\eta(t)$ on some countable state space $E$. Denote by $\mathbf{P}_\xi$, $\xi \in E$, the probability measure on the path space $D(\mathbb{R}_+, E)$ induced by the Markov process $\eta(t)$ starting from $\xi$, and by $\text{cap}(A, B)$ the capacity between two disjoint subsets, $A$, $B$, of $E$. Next result was communicated to us by A. Teixeira.

Lemma 4.3. Let $A$, $B$ be two disjoint subsets of $E$, $A \cap B = \emptyset$, and let $\eta \notin A \cup B$. Then,

$$\mathbf{P}_\eta[H_B < H_A] = \frac{\text{cap}(\eta, A \cup B) + \text{cap}(B, A \cup \{\eta\}) - \text{cap}(A, B \cup \{\eta\})}{2 \text{cap}(\eta, A \cup B)} \leq \frac{\text{cap}(\eta, B)}{\text{cap}(\eta, A \cup B)}.$$

Proof. Denote by $\eta^T(t)$ the trace of the process $\eta(t)$ on the set $A \cup B \cup \{\eta\}$, and by $\mathbf{P}_\eta^T$ the distribution of this Markov process starting from $\eta$. Clearly,

$$\mathbf{P}_\eta[H_B < H_A] = \mathbf{P}_\eta[H_B < H_A] = \frac{R_T(\eta, B)}{R_T(\eta, A) + R_T(\eta, B)},$$

if $R_T(\zeta, \xi)$ represents the jump rates of the trace process $\eta^T(t)$. Denote by $\mu_T(\cdot, \cdot)$ the stationary measure of the process $\eta^T(t)$. Multiplying the numerator and the denominator of the former ratio by $\mu_T(\eta)$, in view of [1 Lemma 6.8], the ratio
Lemma. Therefore, the statement of the lemma may have some interest only if
which is the expression which appears in the denominator in the proof of the
Lemma 4.3. Next result provides a
Corollary 4.4. The second term on the right hand side is the expression we added to the numerator
to transform the identity presented in Lemma 4.3 into an inequality. On the other
hand, since $H_A \wedge H_B < H^+_{\eta}$, the capacity associated to the process $\eta^T(t)$
becomes
\[
\frac{\text{cap}_T(\eta, A \cup B) + \text{cap}_T(B, A \cup \{\eta\}) - \text{cap}_T(A, B \cup \{\eta\})}{2 \text{cap}_T(\eta, A \cup B)}
\]
\[
= \frac{\text{cap}(\eta, A \cup B) + \text{cap}(B, A \cup \{\eta\}) - \text{cap}(A, B \cup \{\eta\})}{2 \text{cap}(\eta, A \cup B)},
\]
where we used [1, Lemma 6.9] in the last equality. This proves the identity. To
derive the inequality, denote by $\lambda_T(\cdot)$ the holding rates of the trace process $\eta^T(t)$
and observe that $\text{cap}_T(\eta, B) = \mu_T(\eta) \lambda_T(\eta) P^T_{\eta}[H_B < H^+_{\eta}] \geq \mu_T(\eta) R_T(\eta, B)$. □

In some cases the estimate presented in the previous lemma has no content. On
the one hand,
\[
\text{cap}_T(\eta, B) = \mu_T(\eta) R_T(\eta, B) + \mu_T(\eta) \lambda_T(\eta) P^T_{\eta}[H_A < H_B < H^+_{\eta}].
\]
The second term on the right hand side is the expression we added to the numerator
to transform the identity presented in Lemma [1, Lemma 6.9] into an inequality. On the other
hand, since $H_A \wedge H_B < H^+_{\eta}$, $P^T_{\eta}$-a.s.,
\[
P^T_{\eta}[H_A < H_B < H^+_{\eta}] = E^T_{\eta}\left[1(H_A < H_B) P^T_{\eta, T(H_A)}[H_B < H^+_{\eta}] \right].
\]
and
\[
\mu_T(\eta) R_T(\eta, B) + \mu_T(\eta) \lambda_T(\eta) P^T_{\eta}[H_A < H_B]
\]
\[
= \mu_T(\eta) R_T(\eta, B) + \mu_T(\eta) R_T(\eta, A) = \text{cap}_T(\eta, A \cup B),
\]
which is the expression which appears in the denominator in the proof of the
lemma. Therefore, the statement of the lemma may have some interest only if
$P^T_{\eta, T(H_A)}[H_B < H^+_{\eta}] = P^T_{\eta, T(H_A)}[H_B < H^+_{\eta}]$ is negligible, i.e., if the process starting
from $A$ reaches $B$ before $\eta$ with a vanishing probability.

We apply Lemma [4.4] to our context to obtain a bound on $P^T_{\eta}[H_{\xi^*_N} \leq \gamma^{-1}]$. For
$\gamma > 0$, consider the Markov process $\{\eta^{N,\gamma}(t) : t \geq 0\}$ on $E_N \cup E^*_N$ whose jump rates
$R_{N,\gamma}(N, \xi) = R_{N,\gamma}(\eta, \xi)$ are given by
\[
R_{N,\gamma}(\eta, \xi) = \begin{cases} 
R_N(\eta, \xi) & \text{if } \eta \text{ and } \xi \in E_N, \\
\gamma & \text{if } \eta \in E^*_N \cup E_N \text{ and if } [\xi = \eta^* \text{ or } \eta = \xi^*], \\
0 & \text{otherwise}. 
\end{cases}
\]
Note that the process $\eta^*(t)$ is the trace of the process $\eta^{N,\gamma}(t)$ on $E^*_N \cup E_N$. Denote
by $\text{cap}_{N,\gamma}$ the capacity associated to the process $\eta^{N,\gamma}(t)$. Next result provides a
bound for condition (L4U) in terms of capacities which can be estimated through
the Dirichlet and the Thomson principles.

**Corollary 4.4.** For every $x \in S$, $\eta \in E^*_N$ and $\gamma > 0$,
\[
P^T_{\eta}[H_{\xi^*_N} \leq \gamma^{-1}] \leq \frac{e \cdot \text{cap}_N(\eta, \xi^*_N)}{2 \text{cap}_{N,\gamma}(\eta, \xi^*_N)}.
\]

**Proof.** By Lemma [4.4] and by Lemma [4.3]
\[
P^T_{\eta}[H_{\xi^*_N} \leq \gamma^{-1}] \leq e \cdot \text{cap}_{\gamma}[H_{\xi^*_N} \leq \gamma^{-1}] \leq \frac{e \cdot \text{cap}(\eta, \xi^*_N)}{\text{cap}_{\gamma}(\eta, \xi^*_N)}.
\]
It is clear from the Dirichlet principle and from the definition of the enlarged process that \(\text{cap}_N(\eta, \tilde{\mathcal{E}}_N) = (1/2)\text{cap}_N(\eta, \tilde{\mathcal{F}}_N)\), where \(\text{cap}_N\) stands for the capacity associated to the trace process \(\eta^N(t)\). By [1, Lemma 6.9], once more, \(\text{cap}_N(\eta, \tilde{\mathcal{E}}_N) = \pi(\mathcal{E}_N)^{-1}\text{cap}_{N}(\eta, \tilde{\mathcal{E}}^*_N)\) and \(\text{cap}_N(\eta, \mathcal{E}^*_N) = \pi(\mathcal{E}_N)^{-1}\text{cap}_{N}(\eta, \mathcal{E}^*_N)\). This concludes the proof of the lemma.

5. Tightness

We prove in this section tightness of the process \(X^N_t\). By Aldous criterion (see Theorem 16.10 in [8]) we just need to show that for every \(\epsilon > 0\) and \(T > 0\)
\[
\lim_{\delta \downarrow 0} \lim_{N \to \infty} \sup_{\tau \in \mathcal{F}_T} \sup_{a \leq \delta} \mathbb{P}^\mathcal{E} \left[ |X^N_{\tau + a} - X^N_{\tau}| > \epsilon \right] = 0 ,
\]
(5.1)
where \(\mathcal{F}_T\) is the set of stopping times bounded by \(T\).

In fact, in the present context of a finite state space, we do not need to consider all stopping times, but just the jump times. More precisely, the process \(X^N_t\) is tight provided
\[
\lim_{\delta \downarrow 0} \lim_{N \to \infty} \sup_{\tau \in \mathcal{F}_T} \sup_{i \geq 0} \mathbb{P}^\mathcal{E} \left[ |X^N_{\tau + 1} - X^N_{\tau}| \leq \delta \right] = 0 ,
\]
where \(\tau_0 = 0\) and \(\tau_i, i \geq 1\), represent the jumping times of the process \(X^N_t\).

Proof of Lemma 2.5. We will prove that (5.1) holds. Fix \(T > 0, \epsilon > 0\) and \(\delta > 0\). By the strong Markov property, for every \(0 < a \leq \delta\) and stopping time \(\tau \leq T\),
\[
\mathbb{P}^\mathcal{E} \left[ |X^N_{\tau + a} - X^N_{\tau}| > \epsilon \right] \leq \sup_{\eta \in \mathcal{E}_N} \mathbb{P}^\mathcal{E}_{\eta} \left[ |X^N_{\tau + a} - X^N_{\tau}| > \epsilon \right] \leq \sup_{\eta \in \mathcal{E}_N} \mathbb{P}^\mathcal{E}_{\eta} \left[ H_{\tilde{\mathcal{E}}_{\tau + a}} \leq \delta \theta_N \right] .
\]
To conclude the proof we need to show that the last term vanishes as \(N \uparrow \infty\) and then \(\delta \downarrow 0\). The arguments used are similar to the ones used in the proof of Theorem 2.4.

Let \(T_N\) be a sequence satisfying the assumptions (L4U). Fix \(x \in S\) and \(\eta \in \mathcal{E}_N\). The probability \(\mathbb{P}^\mathcal{E}_{\eta}[H_{\tilde{\mathcal{E}}_{\tau}} \leq \delta \theta_N]\) is bounded above by
\[
\mathbb{P}^\mathcal{E}_{\eta}[H_{\tilde{\mathcal{E}}_{\tau}} \leq T_N] + \mathbb{P}^\mathcal{E}_{\eta} \left[ \{ H_{\tilde{\mathcal{E}}_{\tau}} > T_N \} \mathbb{P}^\mathcal{E}_{\eta(T_N)}[H_{\tilde{\mathcal{E}}_{\tau}} \leq \delta \theta_N] \right] ,
\]
(5.2)
The first term vanishes in view of assumption (L4U). On the set \(\{ H_{\tilde{\mathcal{E}}_{\tau}} > T_N \}\), we may couple the process \(\eta(t)\) with the reflected process \(\eta^N(t)\) in a way that \(\eta(t) = \eta^N(t)\) for \(0 \leq t \leq T_N\). In particular, we may replace in the previous term \(\mathbb{P}^\mathcal{E}_{\eta(T_N)}[H_{\tilde{\mathcal{E}}_{\tau}} \leq \delta \theta_N]\) by \(\mathbb{P}^\mathcal{E}_{\eta^N(T_N)}[H_{\tilde{\mathcal{E}}_{\tau}} \leq \delta \theta_N]\). After this replacement we may bound the second term in (5.2) by
\[
\sum_{\xi \in \mathcal{E}_N} \left\{ \left( \delta \eta S^\mathcal{E}_{\tau}(T_N) \right)(\xi) - \pi_x(\xi) \right\} \mathbb{P}^\mathcal{E}_{\xi} \left[ H_{\tilde{\mathcal{E}}_{\tau}} \leq \delta \theta_N \right] + \mathbb{P}^\mathcal{E}_{\pi_x} \left[ H_{\tilde{\mathcal{E}}_{\tau}} \leq \delta \theta_N \right] ,
\]
(5.3)
where \(S^\mathcal{E}_{\tau}(t)\) represents the semi-group of the reflected process. The first term of this sum is bounded by \(||\delta \eta S^\mathcal{E}_{\tau}(T_N) - \pi_x||_{\text{TV}}\), where \(||\mu - \nu||_{\text{TV}}\) stands for the total variation distance between \(\mu\) and \(\nu\). By definition of the mixing time, this last expression is less than or equal to \((1/2)(T_N/T_{r_{\text{mix}}})\), which vanishes as \(N \uparrow \infty\) by definition of the sequence \(T_N\).

It remains to estimate the second term in (5.3). It can be written as
\[
\sum_{\eta \in \mathcal{E}_N} \left\{ \pi_x(\eta) - \pi_x(\eta) \right\} \mathbb{P}^\mathcal{E}_{\eta} \left[ H_{\tilde{\mathcal{E}}_{\tau}} \leq \delta \theta_N \right] + \mathbb{P}^\mathcal{E}_{\pi_x} \left[ H_{\tilde{\mathcal{E}}_{\tau}} \leq \delta \theta_N \right] ,
\]
(5.4)
where \( \pi^*_x \) is the quasi-stationary measure associated to the trace process \( \eta^\xi(t) \) killed when it hits \( \tilde{\mathcal{E}}^\xi \). The first term is less than or equal to
\[
\sum_{\eta \in \mathcal{E}_N^\xi} \pi_x(\eta) \left| \frac{\pi_x^*_x(\eta)}{\pi_x(\eta)} - 1 \right| \leq \left\{ \sum_{\eta \in \mathcal{E}_N^\xi} \pi_x(\eta) \left( \frac{\pi_x^*_x(\eta)}{\pi_x(\eta)} - 1 \right)^2 \right\}^{1/2}.
\]

By Proposition 2.1, (17) and Lemma 2.2 in [7], the expression inside the square root on the right hand side is bounded by \( \varepsilon_x/[1 - \varepsilon_x] \), where \( \varepsilon_x = E_{\pi_x}[R^\xi(\eta, \tilde{\mathcal{E}}^\xi)]/\mathcal{G}_{r,x} \). By [22], \( \varepsilon_x \leq C_0(\theta_N \mathcal{G}_{r,x})^{-1} \) for some finite constant \( C_0 \) and by hypothesis, \( \theta_N^{-1} \ll \mathcal{G}_{r,x} \). This shows that the first term in (5.4) vanishes as \( N \uparrow \infty \).

Finally, since \( \pi^*_x \) is the quasi-stationary state, under \( \mathbb{P}_{\pi^*_x} \), the hitting time of \( \tilde{\mathcal{E}}^\xi \), \( H_{\tilde{\mathcal{E}}^\xi} \), has an exponential distribution whose parameter we denote by \( \phi^*_x \). By Lemma 2.2, \( \phi^*_x \) is bounded by \( E_{\pi_x}[R^\xi(\eta, \tilde{\mathcal{E}}^\xi)] \leq C_0/\theta_N \), for some finite constant \( C_0 \). Hence,
\[
\mathbb{P}_{\pi^*_x}[H_{\tilde{\mathcal{E}}^\xi} \leq \delta \theta_N] = 1 - e^{-\phi^*_x \delta \theta_N} \leq 1 - e^{-C_0 \delta},
\]
an expression which vanishes as \( \delta \downarrow 0 \). This proves (5.4) and concludes the proof of the lemma.

By a version of [22, Theorem 12.3] for continuous-time reversible Markov chains, \( T_{r,x}^{\text{mix}} \leq \mathcal{G}_{r,x}^{-1} \log(4/\min_{\eta \in \mathcal{E}_x^\xi} \pi_x(\eta)) \). Hence, \( \max_{x \in \mathcal{S}} T_{r,x}^{\text{mix}} \ll T_N \) if
\[
\lim_{N \to \infty} \frac{1}{T_N \mathcal{G}_{r,x}} \log \left( \frac{1}{\min_{\eta \in \mathcal{E}_x^\xi} \pi_x(\eta)} \right) = 0 . \tag{5.5}
\]

6. THE SPECTRAL GAP OF THE TRACE PROCESS

We prove in this section Proposition 2.1. We start with an elementary result which provides an upper bound for the spectral gap of the trace process in terms of capacities. Recall that \( \eta(t) \) is a positive recurrent, reversible, continuous-time Markov chain on a countable state space \( \mathcal{E}_N \), whose embedded discrete-time chain is also positive recurrent. Let \( \mathcal{E}_N \) a subset of \( \mathcal{E}_N \) and denote by \( \mathcal{G}_\xi \) the spectral gap of the trace of \( \eta(t) \) on \( \mathcal{E}_N \).

**Lemma 6.1.** We have that
\[
\mathcal{G}_\xi \leq \inf_{A \subset \mathcal{E}_N} \frac{\pi(\mathcal{E}_N) \text{cap}(A, \mathcal{B})}{\pi(A) \pi(\mathcal{B})},
\]
where \( \mathcal{B} = \mathcal{E}_N \setminus A \).

**Proof.** Fix a subset \( A \) of \( \mathcal{E}_N \), and let \( \mathcal{B} = \mathcal{E}_N \setminus A \). By definition,
\[
\mathcal{G}_\xi = \inf_{f} \frac{\langle f, (-L_\xi) f \rangle_{\pi_\xi}}{\text{Var}_{\pi_\xi}(f)} \leq \frac{\langle 1\{A\}, (-L_\xi) 1\{A\} \rangle_{\pi_\xi}}{\text{Var}_{\pi_\xi}(1\{A\})},
\]
where \( \text{Var}_{\pi_\xi}(f) \) stands for the variance of \( f \) with respect to the measure \( \pi_\xi \).
Since \( \mathcal{E}_N = A \cup \mathcal{B} \), \( 1\{A\} \) is the equilibrium potential between \( A \) and \( \mathcal{B} \) so that \( \langle 1\{A\}, (-L_\xi) 1\{A\} \rangle_{\pi_\xi} = \text{cap}_\xi(A, \mathcal{B}) \). Hence, by [11, Lemma 6.9],
\[
\mathcal{G}_\xi \leq \frac{\text{cap}_\xi(A, \mathcal{B})}{\pi_\xi(A) \pi_\xi(\mathcal{B})} = \frac{\pi(\mathcal{E}_N) \text{cap}(A, \mathcal{B})}{\pi(A) \pi(\mathcal{B})}. \tag{5.6}
\]
\( \square \)
Proof of Proposition 2.1 Let \( F : \mathcal{E}_N \to \mathbb{R} \) be a function in \( L^2(\pi_\mathcal{E}) \) and denote by \( \hat{F} : E_N \to \mathbb{R} \) the harmonic extension of \( F \) to \( E_N \), defined by

\[
\hat{F}(\eta) = \begin{cases} 
F(\eta) & \text{if } \eta \in \mathcal{E}_N, \\
\mathbb{E}_\eta[F(\eta(H_{E_N}))] & \text{if } \eta \notin \mathcal{E}_N.
\end{cases}
\]

We claim that

\[
\langle (-L_N)\hat{F}, \hat{F} \rangle_\pi = \pi(\mathcal{E}_N) \langle (-L_\mathcal{E})F, F \rangle_{\pi_\mathcal{E}}. 
\]  

(6.1)

Indeed, since \( L_N \hat{F} = 0 \) on \( \mathcal{E}_N \) and since \( \hat{F} \) and \( F \) coincide on \( \mathcal{E}_N \), the Dirichlet form \( \langle L_N \hat{F}, \hat{F} \rangle_\pi \) is equal to

\[
\sum_{\eta \in \mathcal{E}_N} \pi(\eta) F(\eta) \sum_{\xi \in E_N} R_N(\eta, \xi) \{ \hat{F}(\xi) - F(\eta) \}.
\]  

(6.2)

We decompose the previous sum in two expressions, the first one including all terms for which \( \xi \) belongs to \( \mathcal{E}_N \) and the second one including all terms for which \( \xi \) belongs to \( E_N \setminus \mathcal{E}_N \). When \( \xi \) belongs to \( \mathcal{E}_N \), we may replace \( \hat{F} \) by \( F \). The other expression, by definition of \( \hat{F} \) is equal to

\[
\sum_{\eta \in \mathcal{E}_N} \sum_{\xi \in \mathcal{E}_N} \pi(\eta) F(\eta) R_N(\eta, \xi) \sum_{\zeta \in E_N} \mathbb{P}_\xi[H_{E_N} = H_\zeta] \{ F(\zeta) - F(\eta) \}.
\]

Since for \( \eta \in \mathcal{E}_N \),

\[
\mathbb{P}_\eta[H_{E_N}^+ = H_\zeta] = p_N(\eta, \zeta) + \sum_{\xi \notin \mathcal{E}_N} p_N(\eta, \xi) \mathbb{P}_\xi[H_{E_N} = H_\zeta],
\]

and since by [1] Proposition 6.1 \( R^\mathcal{E}(\eta, \zeta) = \lambda_N(\eta) \mathbb{P}_\eta[H_{E_N}^+ = H_\zeta] \) the previous sum is equal to

\[
\sum_{\eta \in \mathcal{E}_N} \sum_{\xi \in \mathcal{E}_N} \pi(\eta) F(\eta) \left\{ R^\mathcal{E}(\eta, \zeta) - R_N(\eta, \zeta) \right\} \{ F(\zeta) - F(\eta) \}.
\]

Adding this sum to the first expression in our decomposition of (6.2) as the sum of two terms, we get that the left hand side of (6.1) is equal to

\[
\sum_{\eta \in \mathcal{E}_N} \sum_{\xi \in \mathcal{E}_N} \pi(\eta) F(\eta) R^\mathcal{E}(\eta, \zeta) \{ F(\zeta) - F(\eta) \}.
\]

To conclude the proof of Claim (6.1), it remains to recall that \( \pi_\mathcal{E}(\eta) = \pi(\eta)/\pi(\mathcal{E}_N) \).

Fix a function \( F : \mathcal{E}_N \to \mathbb{R} \). We claim that

\[
\inf_g \langle (-L_N)g, g \rangle_\pi = \langle (-L_\mathcal{E})\hat{F}, \hat{F} \rangle_\pi, 
\]

(6.3)

where the infimum is carried over all functions \( g : E_N \to \mathbb{R} \) which are equal to \( F \) on \( \mathcal{E}_N \). Indeed, it is simple to show that any function \( f \) which solves the variational problem on the left hand side of (6.3) is harmonic on \( \mathcal{E}_N \) and coincides with \( F \) on \( \mathcal{E}_N \). Indeed, \( L_N f = 0 \) on \( \mathcal{E}_N \) and \( f = F \) on \( E_N \). The unique solution to this problem is \( \hat{F} \), which proves (6.3).

Fix an eigenfunction \( F \) associated to \( \mathcal{E}_N \) such that \( E_{\mathcal{E}_N}[F^2] = 1 \), \( E_{\mathcal{E}_N}[F] = 0 \). By (6.1) we have that

\[
\mathcal{E}_N \langle (-L_\mathcal{E})F, F \rangle_{\pi_\mathcal{E}} = \frac{1}{\pi(\mathcal{E}_N)} \langle (-L_N)\hat{F}, \hat{F} \rangle_\pi.
\]
By the spectral gap, the Dirichlet form on the right hand side is bounded below by \( g \) times the variance of \( \tilde{F} \). This latter variance, in view of the definition of \( \tilde{F} \) and the properties of \( F \), is equal to

\[
\pi(\mathcal{E}_N) + \sum_{\eta \notin \mathcal{E}_N} \pi(\eta)\tilde{F}(\eta)^2 - \left( \sum_{\eta \notin \mathcal{E}_N} \pi(\eta)\tilde{F}(\eta) \right)^2 \geq \pi(\mathcal{E}_N).
\]

This proves that \( \mathcal{G} \leq \mathcal{G}_\varepsilon \).

Fix an eigenfunction \( f \) associated to \( g \) such that \( E_\pi[f^2] = 1, E_\pi[f] = 0 \). Let \( F : \mathcal{E}_N \to \mathbb{R} \) be the restriction to \( \mathcal{E}_N \) of \( f : F(\eta) = f(\eta)1\{\eta \in \mathcal{E}_N\} \). By definition of \( g \),

\[
g = \langle (-L_N)f, f \rangle_\pi \geq \inf_g \langle (-L_N)g, g \rangle_\pi,
\]

where the infimum is carried over all functions \( g \) which coincide with \( F \) on \( \mathcal{E}_N \). By (6.3), by (6.1) and by definition of the spectral gap \( \mathcal{G}_\varepsilon \), the right hand side of the previous term is equal to

\[
\langle (-L_N)\tilde{F}, \tilde{F} \rangle_\pi = \pi(\mathcal{E}_N)\langle (-L_N)F, F \rangle_\pi \geq \mathcal{G}_\varepsilon \pi(\mathcal{E}_N)\left\{ E_\pi[F^2] - E_\pi[F]^2 \right\}.
\]

Since \( F = f1\{\mathcal{E}_N\} \), up to this point we proved that

\[
\mathcal{G}_\varepsilon \pi(\mathcal{E}_N)\left\{ E_\pi[f^21\{\mathcal{E}_N\}] - E_\pi[f1\{\mathcal{E}_N\}]^2 \right\} \leq \mathcal{G}.
\]

Since the eigenfunction \( f \) associated to \( g \) is such that \( E_\pi[f^2] = 1, E_\pi[f] = 0 \), we may rewrite the previous inequality as

\[
\mathcal{G}_\varepsilon \left\{ 1 - \frac{E_\pi[f^21\{\mathcal{E}_N\}]}{\pi(\mathcal{E}_N)} + \frac{1}{\pi(\mathcal{E}_N)}E_\pi[f1\{\mathcal{E}_N\}]^2 \right\} \leq \mathcal{G}.
\]

By Schwarz inequality, \( E_\pi[f1\{\mathcal{E}_N\}]^2 \leq E_\pi[f^21\{\mathcal{E}_N\}] \pi(\mathcal{E}_N) \) so that

\[
\mathcal{G}_\varepsilon \left\{ 1 - \frac{1}{\pi(\mathcal{E}_N)}E_\pi[f^21\{\mathcal{E}_N\}] \right\} \leq \mathcal{G},
\]

which proves the proposition. \( \square \)

7. Proof of Theorem 2.8

We assume in this section that the state space \( E_N \) has been divided in three disjoint sets \( \mathcal{E}_N^A = A, \mathcal{E}_N^B = B \) and \( \Delta_N = E_N \setminus \mathcal{E}_N \), where \( \mathcal{E}_N = A \cup B \). Recall that \( \eta^\varepsilon(t) \) represents the trace of the process \( \eta(t) \) on the set \( \mathcal{E}_N \) and \( \eta^\varepsilon(t) \) the \( \gamma \)-enlargement of the process \( \eta^\varepsilon(t) \) to the set \( \mathcal{E}_N \cup \mathcal{E}_N^\ast \), where \( \gamma = \gamma_N \) is a sequence of positive numbers and \( \mathcal{E}_N^\ast = A^\ast \cup B^\ast, A^\ast, B^\ast \) being copies of the sets \( A, B \), respectively. Denote by \( \mathcal{G}_A, \mathcal{G}_B \) the spectral gap of the process \( \eta(t) \) reflected at \( A, B \), respectively.

Let \( \widehat{\mathcal{C}}p_\pi(A^\ast, B^\ast) \) be the normalized capacity between \( A^\ast \) and \( B^\ast \):

\[
\widehat{\mathcal{C}}p_\pi(A^\ast, B^\ast) = \frac{\mathcal{C}p_\pi(A^\ast, B^\ast)}{\pi_\pi(A) \pi_\pi(B)}.
\]

By [7] Theorem 2.12,

\[
\left( 1 - \frac{2 \widehat{\mathcal{C}}p_\pi(A^\ast, B^\ast)}{\gamma} \right)^2 \leq \frac{2 \widehat{\mathcal{C}}p_\pi(A^\ast, B^\ast)}{\mathcal{G}_\varepsilon} \leq 1 + \frac{\gamma + 2 \widehat{\mathcal{C}}p_\pi(A^\ast, B^\ast)}{\min\{\mathcal{G}_A, \mathcal{G}_B\}} .
\]

(7.1)

The factor 2, which is not present in [7], appears because we consider the capacity with respect to the probability measure \( \pi^\ast \), while [7] defines the capacity with \( \pi_\varepsilon \) as reference measure.
Theorem 2.8 is a simple consequence of (7.1). For sake of completeness, we present a proof of the lower bound of (7.1). Let \( V \) be the equilibrium potential between \( A^* \) and \( B^* \): \( V(\eta) = \mathbb{P}_\eta^\gamma[H_{A^*} < H_{B^*}] \). We sometimes consider below \( V \) as a function on \( \mathcal{E}_N \). By definition of the spectral gap,

\[
\mathfrak{g}_\varepsilon \leq \frac{(V, (-L_\varepsilon)V)_{\pi_\varepsilon}}{\text{Var}_{\pi_\varepsilon}(V)},
\]

where \( \text{Var}_{\pi_\varepsilon}(V) \) stands for the variance of \( V \) with respect to the measure \( \pi_\varepsilon \). We estimate the numerator and the denominator separately.

Since the capacity between \( A^* \) and \( B^* \) is equal to the Dirichlet form of the equilibrium potential,

\[
(1/2)(V, (-L_\varepsilon)V)_{\pi_\varepsilon} \leq \text{cap}_\gamma(A^*, B^*).
\]

A martingale decomposition of the variance of \( V \) gives that

\[
\text{Var}_{\pi_\varepsilon}(V) \geq \pi_\varepsilon(A) \pi_\varepsilon(B) \left( E_{\pi_A}[V_A] - E_{\pi_B}[V_B] \right)^2,
\]

where \( V_A = V1\{A\} \), \( V_B = V1\{B\} \). Since \( \text{cap}_\gamma(A^*, B^*) = (V, (-L_\varepsilon)V)_{\pi_\varepsilon} \), since \( (L_\varepsilon V)(\eta^*) = \gamma(V(\eta) - 1) \), where \( \eta^* \) is the state in \( \mathcal{E}_N^\varepsilon \) corresponding to the state \( \eta \in \mathcal{E}_N \), and since \( \pi_\varepsilon(\eta^*) = (1/2)\pi_\varepsilon(\eta) \), \( 2\text{cap}_\gamma(A^*, B^*) = \gamma \pi_\varepsilon(A) - \gamma \sum_{\eta \in A} \pi_\varepsilon(\eta)V(\eta) \).

Therefore,

\[
E_{\pi_A}[V_A] = 1 - \frac{2\text{cap}_\gamma(A^*, B^*)}{\gamma \pi_\varepsilon(A)}.
\]

Repeating the previous argument with \( 1 - V \) in place of \( V \) we obtain that

\[
E_{\pi_B}[V_B] = \frac{2\text{cap}_\gamma(A^*, B^*)}{\gamma \pi_\varepsilon(B)}.
\]

Putting together the previous estimates, we conclude the proof of the lower bound of (7.1). \qed

8. Applications

We present in this section two applications of Theorems 2.2 and 2.4. Both processes do not visit points in the time scale where tunneling occurs and, therefore, do not satisfy the assumptions of the theory developed in [1] [5]. Furthermore, these models have logarithmic energy or entropy barriers, restraining the application of large deviations methods. On the other hand, both dynamics are monotone with respect to a partial order, allowing the use of coupling techniques. The first model, which has only entropy barriers, was suggested by A. Gaudilliére to the authors as a model for testing metastability techniques. We prove for this model the mixing conditions introduced in Section 2.E. The second one has been examined in details in [12] [11]. We apply to this model the \( L^2 \)-theory presented in Section 2.D.

8.1. The dog graph [24]. For \( N \geq 1 \) and \( d \geq 2 \), let \( Q_N = \{0, \ldots, N\}^d \) be a \( d \)-dimensional cube of length \( N \), let \( \bar{Q}_N \) be the reflection of \( Q_N \) through the origin, \( \bar{Q}_N = \{\eta \in \mathbb{Z}^d : -\eta \in Q_N\} \), and let \( V_N = Q_N \cup \bar{Q}_N \). Denote by \( E_N \) the set of edges formed by pairs of nearest-neighbor sites of \( V_N \), \( E_N = \{(\eta, \xi) \in V_N \times V_N : |\eta - \xi| = 1\} \). The graph \( G_N = (V_N, E_N) \) is called the dog graph [24].

Let \( \{\eta(t) : t \geq 0\} \) be the continuous-time Markov chain on \( G_N \) which jumps from \( \eta \) to \( \xi \) at rate 1 if \( (\eta, \xi) \in E_N \). The uniform measure on \( V_N \), denoted by \( \pi_0 \), is
the unique stationary state. Diaconis and Saloff Coste [24, Example 3.2.5] proved that there exist constants $0 < \frac{c(d)}{N^d} < C(d) < \infty$ such that for all $N \geq 1$,

$$\frac{c(2)}{N^2 \log N} \leq g \leq \frac{C(2)}{N^2 \log N} \quad \text{in } d = 2 \quad \text{and} \quad \frac{c(d)}{N^d} \leq g \leq \frac{C(d)}{N^d} \quad (8.1)$$

in dimension $d \geq 3$.

Fix a sequence $\alpha_N$, $(\log N)^{-1/2} \ll \alpha_N \ll 1$, and let $\mathcal{B}_N = \{ \eta = (\eta_1, \ldots, \eta_d) \in V_N : \min \eta_j \geq \alpha_N N \}$, $A_N = -\mathcal{B}_N = \{ \eta \in V_N : -\eta \in \mathcal{B}_N \}$. Denote by $\rho_{\mathcal{A}}$ and $\rho_{\mathcal{B}}$ (resp. $\rho_{\mathcal{A}, \mathcal{B}}$ and $\rho_{\mathcal{A}, \mathcal{B}}^{\text{mix}}$) the spectral gap and the mixing time of the continuous-time random walk $\eta(t)$ reflected at $A_N$ (resp. $\mathcal{B}_N$). It is well known that there exist finite constants $0 < c(d) < C(d) < \infty$ such that for all $N \geq 1$,

$$\frac{c(d)}{N^2} \leq \rho_{\mathcal{A}} \leq \frac{C(d)}{N^2}, \quad c(d) N^2 \leq \rho_{\mathcal{A}, \mathcal{B}}^{\text{mix}} \leq C(d) N^2, \quad (8.2)$$

with similar inequalities if $\mathcal{B}$ replaces $\mathcal{A}$.

**Condition (2.10).** Let $E_N = A_N \cup \mathcal{B}_N$, and recall the notation introduced in Section 2. We claim that condition (2.10) is fulfilled for $\theta_N = N^2 \log N$ in dimension 2 and for $\theta_N = N^d$ in dimension $d \geq 3$. Indeed, if $\pi_{\mathcal{A}}$, $\pi_{\mathcal{B}}$, $\pi_{\mathcal{E}}$ represent the uniform measure $\pi$ conditioned to $A_N$, $\mathcal{B}_N$, $E_N$, respectively, by (2.11),

$$E_{\pi_{\mathcal{A}}} \left[ R^E (\eta, \mathcal{B}) \right] = \frac{1}{\pi(A_N)} \text{cap}_N (A_N, \mathcal{B}_N).$$

By the Dirichlet principle, the capacity is bounded by the Dirichlet form of any function which vanishes on $A_N$ and is equal to 1 on $\mathcal{B}_N$. In dimension $d \geq 3$ we simply choose the indicator of the set $Q_N$. In dimension 2, let $D_k = \{ \eta \in Q_N : \eta_1 + \eta_2 = k \}$, $k \geq 0$. Fix $1 \leq L \leq N$ and consider the function $f_L : Q_N \to \mathbb{R}_+$ defined by $f(0) = 0$,

$$f_L(\eta) = \frac{1}{\Phi(L)} \sum_{j=1}^k \frac{1}{j} \eta \in D_k, \quad 1 \leq k \leq L, \quad (8.3)$$

where $\Phi(L) = \sum_{1 \leq j \leq L} j^{-1}$, and $f_L(\eta) = 1$ otherwise. It is easy to see that the Dirichlet form of $f_L$ is bounded by $C_0 (N^2 \log L)^{-1}$ for some finite constant $C_0$. Choosing $L = N^{1/2}$, we conclude that there exists a finite constant $C_0$ such that

$$\text{cap}_N(A_N, \mathcal{B}_N) \leq C_0 \quad \text{for } d \geq 2, \quad \text{cap}_N(A_N, \mathcal{B}_N) \leq C_0 N^d, \quad d \geq 3. \quad (8.4)$$

Condition (2.10) follows from this estimate and the definition of the sequence $\theta_N$. **Condition (L1B) in Lemma 2.9** By Lemma 6.1 and by the previous estimate of the capacity, there exists a finite constant $C_0$ such that $g_{\mathcal{E}} \leq C_0 [N^2 \log N]^{-1}$ in dimension 2 and $g_{\mathcal{E}} \leq C_0 N^{-d}$ in dimension $d \geq 3$. Condition (L1B) is thus fulfilled in view of (8.2).

**Condition (L4) in Theorem 2.4** We claim that there exists a sequence $T_N$ satisfying the conditions (L4) if $\nu_N$ is a sequence of measures concentrated on $A_N$ and such that

$$\lim_{N \to \infty} \frac{1}{R_N} E_{\pi_{\mathcal{E}}} \left[ \left( \frac{\nu_N}{\pi_{\mathcal{E}}} \right)^2 \right] = 0, \quad (8.5)$$

where $R_N = \log N$ in dimension $d = 2$, and $R_N = N^{d-2}$ in dimension $d \geq 3$. Let $M_N$ be an increasing sequence, $M_N \gg 1$, for which (8.5) still holds if multiplied by
Consider the case of dimension 2. In view of the definition of \( T \) and (Condition T) of Lemma 8.1, it is enough to show that \( P_{\nu_N}[H_{\mathcal{B}_N} \leq N^2 M_N] = 0 \).

The strategy proposed in Section 6 permits to weaken assumption (8.5).

**Lemma 8.1.** Let \( T_N \) be a sequence such that \( T_N \ll \alpha_N^2 N^2 \log N \) in dimension 2, and \( T_N \ll \alpha_N^d N^d \) in dimension \( d \geq 3 \). Then,

\[
\lim_{N \to \infty} \max_{\eta \in \mathcal{A}_N} P_{\eta}^* [H_{\mathcal{B}_N} \leq T_N] = 0.
\]

**Proof.** In view of the definition of \( \alpha_N \), we may assume that \( T_N \gg N^2 \). We present the arguments in dimension 2, the case of higher dimension being similar. Fix a sequence \( \eta^N \in \mathcal{A}_N \). Let \( \gamma_N = T_N^{-1} \) and denote by \( \eta^*(t) \) the \( \gamma \)-enlargement of the process \( \eta(t) \) on \( V_N \cup V_N^* \), as defined in Section 2. Here, \( V_N^* \) represents a copy of \( V_N \), and the process \( \eta^*(t) \) jumps from \( \eta \) to \( \eta^* \) (and from \( \eta^* \) to \( \eta \)) at rate \( \gamma_N \). Denote by \( P_{\eta}^* \) the probability measure on the path space \( D(\mathbb{R}_+, V_N \cup V_N^*) \) induced by the Markov process \( \eta^*(t) \) starting from \( \eta \).

Let \( W \) be the equilibrium potential \( W(\eta) = P_{\eta}^*[H_0 < H_{\tilde{Q}_N}^*] \), where 0 represents the origin. In view of Lemma 4.1, it is enough to show that \( W(\eta^N) \) vanishes as \( N \uparrow \infty \). By the Dirichlet principle,

\[
(W, (-L_*)W)_{\pi_*} = \text{cap}_*(0, \tilde{Q}_N^*) = \inf_{f} \langle f, (-L_*)f \rangle_{\pi_*},
\]

where the infimum is carried over all functions \( f \) which vanish on \( \tilde{Q}_N^* \) and which are equal to 1 at the origin. Using the function \( f_L \) introduced in (8.3), we may show that the last term is bounded by \( C_0 (N^2 \log N)^{-1} \) for some finite constant \( C_0 \). We used here the fact that \( \gamma_N \ll (N \log N)^{-1} \).

Denote by \( \prec \) the partial order of \( \mathbb{Z}^d \) so that \( \eta \prec \xi \) if \( \eta_j \leq \xi_j \) for \( 1 \leq j \leq d \). A coupling argument shows that the equilibrium potential \( W \) is monotone on \( Q_N^* \): \( W(\eta) \leq W(\xi) \) for \( \eta \prec \xi \), \( \eta, \xi \in Q_N^* \). Suppose that \( W(\eta^N) \) does not vanish as \( N \uparrow \infty \). In this case there exists \( \epsilon > 0 \) and a subsequence \( N_j \), still denoted by \( N \), such that \( W(\eta^N) \geq \epsilon \) for all \( N \). Let \( U_N = \{ \xi \in Q_N^*: \eta^N \prec \xi \} \). By monotonicity of the equilibrium potential, \( W(\xi) \geq W(\eta^N) \geq \epsilon \) for all \( \xi \in U_N \). Therefore,

\[
(W, (-L_*)W)_{\pi_*} \geq \gamma_N \sum_{\xi \in U_N} \pi_*(\xi) W(\xi)^2 \geq c_0 \gamma_N \epsilon^2 \alpha_N^2
\]

for some positive constant \( c_0 \). This contradicts the estimate (8.6) because \( \gamma_N \gg (\alpha_N^2 N^2 \log N)^{-1} \).

**Condition (L4U).** The proof of Lemma 8.1 shows that condition (L4U) is in force.

**Lemma 8.2.** Let \( T_N \) be a sequence such that \( T_N \ll \alpha_N^2 N^2 \log N \) in dimension 2, and \( T_N \ll \alpha_N^d N^d \) in dimension \( d \geq 3 \). Then,

\[
\lim_{N \to \infty} \max_{\eta \in \mathcal{A}_N} P_{\eta}^*[H_{\mathcal{B}_N} \leq T_N] = 0.
\]

**Proof.** Consider the case of dimension 2. In view of the definition of \( \alpha_N \), we may assume that \( T_N \gg N^2 \). Let \( \eta_N = T_N^{-1} \), and fix a sequence \( \eta^N \in \mathcal{A}_N \). By the proof of Lemma 8.1 it is enough to show that \( P_{\eta^N}^*[H_{\mathcal{B}_N} < H_{\mathcal{A}_N^*}] \) vanishes as \( N \uparrow \infty \),
Hence, vanishes as Lemma 2.6 are fulfilled for $N$. We have shown in the proof of Lemma 8.1 that the first term on the right hand side of the previous formula vanishes as $N \to \infty$. It is therefore enough to show that

$$\lim_{N \to \infty} \max_{\eta \in \mathcal{Q}_N} W(\eta) = 0.$$ 

Let $R_N$ be a sequence such that $N^2 \ll R_N \ll T_N$. With respect to the measure $\mathbb{P}_r^{\mathbf{Q}^\gamma} \mathbf{H}_{\mathbf{Q}_N^*}$, a mean $T_N$ exponential time. Hence, $\mathbb{P}_r^{\mathbf{Q}^\gamma}[H_{\mathbf{Q}_N^*} < R_N]$ vanishes as $N \to \infty$. It is therefore enough to show that

$$\lim_{N \to \infty} \mathbb{P}_r^{\mathbf{Q}^\gamma}[H_{\mathbf{A}_N^1, \mathbf{Q}_N^*} < H_{\mathbf{A}_N^2}, H_{\mathbf{Q}_N^*} \geq R_N] = 0.$$ 

By the Markov property, the previous probability is equal to

$$\mathbb{P}_r^{\mathbf{Q}^\gamma}[1\{H_{\mathbf{Q}_N^*} \geq R_N\} \mathbb{P}_r^{\mathbf{Q}^\gamma}[H_{\mathbf{A}_N^1} < H_{\mathbf{A}_N^2}]],$$

where $\mathbf{Q}^\gamma(t)$ is the process $\eta(t)$ reflected at $\mathbf{Q}_N$. We bound last expectation by removing the indicator of the set $H_{\mathbf{Q}_N} \geq R_N$ and we estimate the remaining term by

$$\mathbb{P}_r^{\mathbf{Q}^\gamma}[H_{\mathbf{A}_N^1} < \mathbf{H}_{\mathbf{A}_N^2}] + \|\delta_r \mathbf{Q}^\gamma(R_N) - \pi_{\mathbf{Q}}\|_{VT},$$

where $\pi_{\mathbf{Q}}$ is the uniform measure on $\mathbf{Q}$ and $\mathbf{Q}^\gamma(t)$ the Markov semigroup of the process $\eta^\mathbf{Q}(t)$. As $R_N \gg N^2$, which is the mixing time of $\eta^\mathbf{Q}(t)$, the second term vanishes as $N \to \infty$, while the first term is the expectation of the equilibrium potential $W$ with respect to the measure $\pi_{\mathbf{Q}}$. If $L_{\mathbf{Q},\mathbf{Q}}$ represents the generator of the Markov process $\eta^\mathbf{Q}(t)$, we have that $L_{\mathbf{Q},\mathbf{Q}} W - \gamma W = -\gamma 1\{A_{N}^1\}$. Taking the expectation with respect to $\pi_{\mathbf{Q}}$, we conclude that $E_{\pi_{\mathbf{Q}}}[W] = \pi_{\mathbf{Q}}(A_{N}^1)$, which vanishes as $N \to \infty$. This concludes the proof of the lemma. 

In view of Lemma 2.10 we have just shown that all assumptions of Theorem 2.4 and Lemma 2.5 are in force. Moreover, by [S2] and Lemma 3.1 the hypotheses of Lemma 2.6 are fulfilled for $\mathcal{D}_N = \mathcal{A}_N$, $\mathcal{F}_N = \mathcal{Q}_N$ and $N^2 \ll T_N \ll \alpha_N^2 (\log N)^2 N^2$. Hence,
Proposition 8.3. Consider the Markov process \( \eta(t) \) on the dog graph. Assume that the initial state \( \nu_N \) is concentrated on \( A_N \). Then, the time-rescaled order \( X^N_t = X^N_{\theta^{-1}_t} \) converges to the Markov process on \( \{1, 2\} \) which starts from 1 and jumps from \( x \) to \( 3-x \) at rate \( 1/2 \). Moreover, in the time scale \( \theta^{-1}_t \) the time spent by the original process \( \eta(t) \) on the set \( \Delta_N = V_N \setminus \mathcal{E}_N \) is negligible.

As a last step, we replace in the previous statement the spectral gap \( g_\mathcal{E} \) of the trace process by the spectral gap \( g \) of the original process. Let \( T_N \) be a sequence such that \( N^2 \ll T_N \ll N^2 \log N \) in dimension 2, and \( N^2 \ll T_N \ll N^d \) in dimension \( d \geq 3 \). It follows from Lemma [8,1] and from [8,2] that

\[
\lim_{N \to \infty} \min_{\eta \in A_N} \mathbb{P}_\eta[\eta(T_N) \in \bar{Q}_N] = 1,
\]

\[
\lim_{N \to \infty} \max_{\eta \in A_N} \| \delta_\eta S(T_N) - \pi_{\bar{Q}_N} \|_{TV} = 0,
\]

where \( S(t) \) is the semigroup of the Markov process \( \eta(t) \). These estimates are the two ingredients needed in the proof of [11] Proposition 2.9, a result which states, among other things, that there exists a mean zero eigenfunction \( f_N \) of the generator \( L_N \) associated to the eigenvalue \( g \) such that \( \lim_N \| f_N \|_\infty = 1 \). By this result and Proposition 2.1

\[
\lim_{N \to \infty} g(E(g/\mathcal{E})) = 1.
\]

8.2. A polymer in the depinned phase [12,11]. 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 consider the dynamics on \( E_N \) induced by the generator \( L_N \) defined by

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

for every function \( f : E_N \to \mathbb{R} \). In this formula \( \eta^{j,\pm} \) represents the configuration which is equal to \( \eta \) 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 \( \eta \) 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 configuration which fulfill the condition \( \eta_{j-1} = \eta_{j+1} = \eta_j + 1 \). Let \( -\eta \) stand for the configuration \( \eta \) reflected at the origin, \((-\eta)_j = -\eta_j, -N \leq j \leq N \). 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 \( \eta \), \( \Sigma(\eta) = \sum_{-N \leq j \leq N} 1 \{ \eta_j = 0 \} \). The probability measure \( \pi_N \) on \( E_N \) defined by \( \pi_N(\eta) = Z^{-1}_{2N} \alpha^{\Sigma(\eta)} \), where \( Z_{2N} \) is a normalizing constant, is easily seen to be reversible for the dynamics generated by \( L_N \).

By [12] Theorem 3.5, the spectral gap \( g \) is bounded above by \( C(\alpha)(\log N)^{8}/N^{5/2} \) for some finite constant \( C(\alpha) \). Following [11], let \( \mathcal{E}_N^1 \) be the set of configurations in \( E_N \) such that \( \eta_j > 0 \) for all \( -(N - \ell) < j < (N - \ell) \), where \( \ell = \ell_N \) is a sequence such that \( 1 \ll \ell_N \ll N \), and let \( \mathcal{E}_N^2 = \{ \eta \in E_N : \eta \notin \mathcal{E}_N^1 \} \), \( \Delta_N = E_N \setminus (\mathcal{E}_N^1 \cup \mathcal{E}_N^2) \).
By equation (2.27) in [11], \( \pi(\mathcal{E}_N) = \pi(\mathcal{E}_N) = (1/2) + O(\ell^{-1/2}) \). Moreover, taking \( \ell_N = (\log N)^{1/4} \), by [11] Proposition 2.6, for every \( \epsilon > 0 \), there exists \( N_0 \) such that for all \( N \geq N_0 \), \( g_{r,1} = g_{r,2} \geq N^{-2+\epsilon} \). In conclusion, choosing \( \epsilon \) small enough and \( \ell_N = (\log N)^{1/4} \),

\[
\|g\| \ll \min \{g_{r,1}, g_{r,2}\}
\]

for all \( N \) large enough, which proves that condition (L1B) is in force.

By [11] Proposition 2.9, there exists an eigenfunction \( f \) of the generator \( L_N \) such that \( E_\pi[f] = 0 \), \( E_\pi[f^2] = 1 \), \( L_N f = \mathcal{g} f \) and \( \|f\|_\infty = 1 + o_N(1) \) where \( \|f\|_\infty \) represents the sup norm of \( f \) and \( o_N(1) \) an expression which vanishes as \( N \uparrow \infty \). Therefore, since \( \pi(\Delta_N) \to 0 \), by Proposition 2.1 \( \mathcal{g}/\mathcal{g}_\ell \) converges to 1 as \( N \uparrow \infty \).

Let \( \nu_N \) be a sequence of probability measures concentrated on \( \mathcal{E}_N \) satisfying condition (2.16). For example, one may define \( \nu_N(\cdot) \) as \( \pi(\cdot | \mathcal{F}) \), where \( \mathcal{F} \) is a subset of \( \mathcal{E}_N \) such that \( \liminf_{N \to \infty} \pi(\mathcal{F}) \geq c_0 \) for some positive constant \( c_0 \). Define the trace process \( \eta^N(t) \) and the order \( X^N_t \) as in Section 2. By Proposition 2.1 and Lemma 2.9 and in view of the previous remarks, the time-rescaled process \( X^N_t = X^N_t/\mathcal{g} \) converges to a Markov process on \( \{1, 2\} \) which starts from 1 and jumps from \( x \) to \( 3-x \) at rate \( 1/2 \). Moreover, by Lemma 2.3, the time spent by the process \( \eta(t) \) on the time scale \( \mathcal{g}^{-1} \) outside the set \( \mathcal{E}_N \) is negligible.

The difference between this result, derived from a general statement, and Theorems 1.3 and 1.5 in [11] is that we require in Theorem 2.2 the initial state to be close to the stationary state of the reflected process in one of the wells, while [11] allows the process to start from any state in one of the wells. This strong assumption on the initial condition permits to consider larger wells and to have an explicit description of these wells. To prove tunneling for a process starting from a state, one needs to show that the mixing conditions (L4U) are in force.

Acknowledgments. The authors wish to thank A. Gaudilli`ere, M. Jara, H. Lacoin, M. Loulakis and A. Teixeira for stimulating discussions.

References

[1] J. Beltrán, C. Landim: Tunneling and metastability of continuous time Markov chains. J. Stat. Phys. 140, 1065–1114, (2010).
[2] J. Beltrán, C. Landim: Metastability of reversible condensed zero range processes on a finite set. Probab. Th. Relat. Fields 152 781–807 (2012)
[3] J. Beltrán, C. Landim: Metastability of reversible finite state Markov processes. Stoch. Proc. Appl. 121 1633–1677 (2011).
[4] J. Beltrán, C. Landim: Tunneling of the Kawasaki dynamics at low temperatures in two dimensions. To appear in Ann. Inst. H. Poincar, Probab. Statist. arXiv:1105.2776 (2011).
[5] J. Beltrán, C. Landim: Tunneling and metastability of continuous time Markov chains II. J. Stat. Phys., , (2012).
[6] A. Bianchi, A. Bovier, D. Ioffe: Pointwise estimates and exponential laws in metastable systems via coupling methods. Ann. Probab. 40 339–371 (2012).
[7] A. Bianchi, A. Gaudilli`ere; Metastable states, quasi-stationary and soft measures, mixing time asymptotics via variational principles. arXiv:1103.1143 (2011)
[8] P. Billingsley, Convergence of Probability Measures, 2nd Edition. John Wiley & Sons, 1999.
[9] 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)
[10] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein. Metastability and low lying spectra in reversible Markov chains. Commun. Math. Phys. 228, 219–255 (2002).
[11] P. Caputo, H. Lacoin, F. Martinelli, F. Simenhau and F. L. Toninelli: Polymer dynamics in the depinned phase: metastability with logarithmic barriers. Probab. Theory Related Fields 153, 587–641 (2012).
[12] 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).
[13] M. Cassandro, A. Galves, E. Olivieri, M. E. Vares. Metastable behavior of stochastic dynamics: A pathwise approach. J. Stat. Phys. 35, 603–634 (1984).
[14] A. Gaudilliè re. Condenser physics applied to Markov chains: A brief introduction to potential theory. Online available at [http://arxiv.org/abs/0901.3653](http://arxiv.org/abs/0901.3653).
[15] A. Gaudilliè re, C. Landim: A Dirichlet principle for non reversible Markov chains and some recurrence theorems. [arXiv:1111.2445](http://arxiv.org/abs/1111.2445) (2011).
[16] B. Gois, C. Landim: Zero-temperature limit of the Kawasaki dynamics for the Ising lattice gas in a large two-dimensional torus. [arXiv:1305.4542](http://arxiv.org/abs/1305.4542) (2013).
[17] M. Jara, C. Landim, A. Teixeira; Quenched scaling limits of trap models. Ann. Probab. 39, 176–223 (2011).
[18] M. Jara, C. Landim, A. Teixeira; Universality of trap models in the ergodic time scale. Online available at [http://arxiv.org/abs/1208.5675](http://arxiv.org/abs/1208.5675) (2012).
[19] 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.
[20] H. Lacoin, A. Teixeira; private communication.
[21] C. Landim: Metastability for a non-reversible dynamics: the evolution of the condensate in totally asymmetric zero range processes. [arXiv:1204.5987](http://arxiv.org/abs/1204.5987).
[22] D. A. Levin, Y. Peres, and E. L. Wilmer. Markov chains and mixing times. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.
[23] E. Olivieri and M. E. Vares. Large deviations and metastability. Encyclopedia of Mathematics and its Applications, vol. 100. Cambridge University Press, Cambridge, 2005.
[24] L. Saloff-Coste. Lectures on finite Markov chains. Lectures on probability theory and statistics (Saint-Flour, 1996), 301–413, Lecture Notes in Math., 1665, Springer, Berlin, 1997.

IMCA, CALLE LOS BÍOLOGOS 245, URB. SAN CÉSAR PRIMERA ETAPA, LIMA 12, PERÚ AND PUCP, AV. UNIVERSITARIA CIDRA. 18, SAN MIGUEL, AP. 1761, LIMA 100, PERÚ.
e-mail: johel.beltran@pucp.edu.pe

IMPA, ESTRADA DONA CASTORINA 110, CEP 22460 RIO DE JANEIRO, BRASIL, AND CNRS UMR 6085, UNIVERSITÉ DE ROUEN, AVENUE DE L’UNIVERSITÉ, BP.12, TECHNOPOLE DU MADRILLET, F76801 SAINT-ÉTIENNE-DU-ROUVRAY, FRANCE.
e-mail: landim@impa.br