ELSEVIER

Metastability of reversible finite state Markov processes[☆]

J. Beltrán[a], C. Landim[b,c,*]

[a] IMCA, Calle los Biólogos 245, Urb. San César Primera Etapa, Lima 12, Perú and PUCP, Av. Universitaria cdra. 18, San Miguel, Ap. 1761, Lima 100, Peru
[b] IMPA, Estrada Dona Castorina 110, CEP 22460 Rio de Janeiro, Brazil
[c] CNRS UMR 6085, Université de Rouen, Avenue de l'Université, BP 12, Technopôle du Madrillet, F76801 Saint-Étienne-du-Rouvray, France

Abstract

We prove the metastable behavior of reversible Markov processes on finite state spaces under minimal conditions on the jump rates. To illustrate the result we deduce the metastable behavior of the Ising model with a small magnetic field at very low temperature.
ⓒ 2011 Elsevier B.V. All rights reserved.

Keywords: Metastability; Finite state Markov processes; Reversibility

1. Introduction

Metastability is a phenomenon observed in thermodynamic systems close to a first-order phase transition. Describing the evolution among competing metastable states or from a metastable state to a stable state in stochastic lattice spin systems at low temperatures is still a subject of considerable interest. We refer the reader to [9,23,4,10] for recent monographs on the subject.

Inspired by the metastable behavior of condensed zero-range processes [1] and from the metastable behavior of random walks among random traps [15,16], we proposed in [2] a

---

definition of metastability and developed some techniques, particularly effective in the reversible case, for proving the metastability of sequences of Markov processes on countable state spaces.

To present the approach introduced in [2] in the simplest possible context, we examine in this article the metastable behavior of reversible Markov processes on finite state spaces. The main result of the article, Theorem 2.1, describes all metastable behaviors of such processes in all time scales under the minimal conditions (2.1), (2.2) on the jump rates.

The minimal assumptions (2.1), (2.2) are clearly satisfied by all Markovian dynamics studied so far. This includes the Glauber dynamics with a small external field at very low temperature in two [20,21] and three [3] dimensions, anisotropic Glauber dynamics [17,18], conservative Kawasaki dynamics [11–14], Markov processes with exponential transition probabilities [24] and the reversible dynamics considered in [7].

Theorem 2.1 asserts the existence of time scales in which a metastable behavior is observed. To apply this result to specific models, as pointed out in Remark 2.2, one needs to compute the capacity between metastable sets and the hitting probabilities of metastable sets. In some cases, as in the Kawasaki dynamics, the exact calculation of the hitting probabilities is impossible, but one can at least determine whether the asymptotic hitting probability is strictly positive or not. In these cases, an exact description of the metastable behavior of the process is not available. It is only known that asymptotically the process spends an exponential time, of a computable mean, in a metastable set, at the end of which it jumps to some other metastable set with an unknown probability, where the same phenomenon is observed.

In contrast with the pathwise approach proposed in [8], the one presented in this article does not highlight the saddle configurations visited by the process when moving from one metastable state to another. However, to compute the exact depths of the valleys, a calculation which relies on a precise estimation of the capacities, one needs to characterize the saddle configurations. This is clearly illustrated in Section 8 where the saddle configurations of a valley $\mathcal{E}_\sigma$, denoted by $\mathbb{W}(\sigma)$, appear when we compute the capacities between the metastable sets of the Ising model.

The lack of precise results on the saddle configurations is compensated by an exact description of the asymptotic dynamics among wells. We are able, in particular, with methods similar to the ones introduced in Bovier et al. [5,6], to show the existence of sequences θ_N for which T_N/θ_N converges to a mean 1 exponential distribution, if T_N represents the time at which the process leaves a metastable set. Furthermore, we also prove the asymptotic independence of T_N/θ_N and η_{T_N}, where η represents the Markov process, a question not considered before. The proof of this asymptotic independence requires the convergence of the average jump rates, defined in (2.6), which is, in most cases, the main technical difficulty in the deduction of metastability.

To illustrate the main result, we consider in Section 3 the metastable behavior of the two-dimensional Ising model with a small external field at very low temperature, the model of Neves and Schonmann [20,21], and a case in which all parameters can be exactly computed.

2. Notation and results

We say that a sequence of positive real numbers $(\alpha_N : N \geq 1)$ is of lower magnitude than a similar sequence $(\beta_N : N \geq 1)$, $\alpha_N \prec \beta_N$ or $\beta_N \succ \alpha_N$, if α_N/β_N vanishes as $N \uparrow \infty$. We say that two positive sequences $(\alpha_N : N \geq 1)$, $(\beta_N : N \geq 1)$ are of the same magnitude, $\alpha_N \approx \beta_N$, if there exists a finite positive constant C_0 such that

$$C_0^{-1} \leq \liminf_{N\to\infty} \frac{\alpha_N}{\beta_N} \leq \limsup_{N\to\infty} \frac{\alpha_N}{\beta_N} \leq C_0.$$

Finally, $\alpha_N \preceq \beta_N$ or $\beta_N \succeq \alpha_N$ means that $\alpha_N \prec \beta_N$ or $\alpha_N \approx \beta_N$.

We say that a set of sequences $(\alpha_N(1) : N \geq 1), \ldots, (\alpha_N(\ell) : N \geq 1)$ is *comparable* if for each $i \neq j$ one of three possibilities holds: $\alpha_N(i) \prec \alpha_N(j)$ or $\alpha_N(j) \prec \alpha_N(i)$ or $\alpha_N(i)/\alpha_N(j)$ converges to a constant $c_{i,j} \in (0, \infty)$. Hence, for example, the possibility that the sequence $\alpha_N(i)/\alpha_N(j)$ oscillates between two finite values and does not converge is excluded.

Fix a finite set E and sequences $\{\lambda_N(j) : N \geq 1\}, 0 \leq j \leq \mathfrak{n}$, such that $\lambda_N(\mathfrak{n}) \prec \lambda_N(\mathfrak{n}-1) \prec \cdots \prec \lambda_N(0) \equiv 1$. Consider a Markov process $\{\eta_t^N : t \geq 0\}$ on E with jump rates denoted by $R_N(x, y), x \neq y \in E$. We assume that the process is *irreducible*, that the unique stationary state, denoted by μ_N, is *reversible*, and that the jump rates satisfy the following multi-scale hypothesis. There exists $a : E \times E \to \{0, \ldots, \mathfrak{n}\}$ such that

$$R_N(x, y) = r_N(x, y) \, \lambda_N(a(x, y)), \quad x \neq y \in E, \tag{2.1}$$

where $\lim_{N \to \infty} r_N(x, y) = r(x, y) \in (0, \infty), x \neq y$. We assume, without loss of generality, that $a(x, y) = 0$ for some $x \neq y$. We assume, furthermore, that products of the rates $\lambda_N(j)$ are comparable. More precisely, we suppose that for any $(k_1, \ldots, k_\mathfrak{n}) \in \mathbb{Z}^\mathfrak{n}$,

$$\lim_{N \to \infty} \prod_{i=1}^{\mathfrak{n}} \lambda_N(i)^{k_i} = C_0 \tag{2.2}$$

for some constant $C_0 \in [0, \infty]$ which depends on $(k_0, \ldots, k_\mathfrak{n})$.

Fix two states x, y in E. By irreducibility, there exists a path $x = x_0, x_1, \ldots, x_n = y$ such that $n \leq |E|, R_N(x_i, x_{i+1}) > 0, 0 \leq i < n$. By the detailed balance condition,

$$\frac{\mu_N(y)}{\mu_N(x)} = \prod_{i=0}^{n-1} \frac{R_N(x_i, x_{i+1})}{R_N(x_{i+1}, x_i)}. \tag{2.3}$$

It follows from assumptions (2.1) and (2.2) that the sequences $\{\mu_N(x) : N \geq 1\}, x \in E$, are comparable. In fact, there exist $\mathfrak{m} \geq 1$, sequences $M_N(\mathfrak{m}) \prec \cdots \prec M_N(1) \prec M_N(0) \equiv 1$, a function $b : E \to \{0, \ldots, \mathfrak{m}\}$ and a function $m : E \to (0, \infty)$ such that

$$\mu_N(x) = m_N(x) \, M_N(b(x)), \quad x \in E, \tag{2.4}$$

where $\lim_{N \to \infty} m_N(x) = m(x) \in (0, \infty)$. We may choose each sequence $M_N(j)$ to be equal to $\prod_{i=1}^{\mathfrak{n}} \lambda_N(i)^{k_i}$ for an appropriate choice of $(k_1, \ldots, k_\mathfrak{n})$ with $\sum_i |k_i| \leq 4|E|$.

Let $G_N : E \times E \to \mathbb{R}_+$ be given by $G_N(x, y) = \mu_N(x)R_N(x, y)$ and note that G_N is symmetric. As above, by (2.1) and (2.2) the sequences $\{G_N(x, y) : N \geq 1\}, x \neq y \in E$, are comparable. Moreover, there exist $\mathfrak{j} \geq 1$, sequences $G_N(\mathfrak{j}) \prec \cdots \prec G_N(1) \prec G_N(0) \equiv 1$, a function $c : E \times E \to \{0, \ldots, \mathfrak{j}\}$ and a function $g : E \to (0, \infty)$ such that

$$G_N(x, y) = g_N(x, y) \, G_N(c(x, y)), \quad x, y \in E, \tag{2.5}$$

where $\lim_{N \to \infty} g_N(x, y) = g(x, y) \in (0, \infty)$. Here also each sequence $G_N(j)$ may be chosen equal to $\prod_{i=1}^{\mathfrak{n}} \lambda_N(i)^{k_i}$ for an appropriate choice of $(k_1, \ldots, k_\mathfrak{n})$ with $\sum_i |k_i| \leq 4|E| + 1$.

Denote by $\mathbf{P}_x^N, x \in E$, the probability measure on the path space $D(\mathbb{R}_+, E)$ induced by the Markov process $\{\eta_t^N : t \geq 0\}$ starting from x. Expectation with respect to $\mathbf{P}_x^N$ is denoted by $\mathbf{E}_x^N$.

For a subset A of E, denote by T_A the time of hitting of A:

$$T_A = \inf\{t > 0 : \eta_t^N \in A\}.$$

When A is a singleton $\{x\}$, we denote $T_{\{x\}}$ by T_x.

For a proper subset F of E, denote by $\{\eta_t^F : t \geq 0\}$ the trace of the Markov process $\{\eta_t^N : t \geq 0\}$ on F. We refer the reader to [2, Section 2] for a precise definition. η_t^F is a Markov process on F and we denote by $R_N^F(x, y)$, $x, y \in F$, its jump rates. Let $r_N^F(A, B)$, $A, B \subset F$, $A \cap B = \varnothing$, be the average rates of jumping of η_t^F from A to B:

$$r_N^F(A, B) = \frac{1}{\mu_N(A)} \sum_{x \in A} \mu_N(x) \sum_{y \in B} R_N^F(x, y). \tag{2.6}$$

The main theorem of this article describes all metastable behaviors of the process $\{\eta_t^N : t \geq 0\}$.

Theorem 2.1. *There exist $\mathfrak{M} \geq 1$, sequences $\{\theta_N(k) : N \geq 1\}$, $1 \leq k \leq \mathfrak{M}$, $1 \prec \theta_N(1) \prec \cdots \prec \theta_N(\mathfrak{M})$, and, for each $1 \leq k \leq \mathfrak{M}$, a partition $\mathcal{E}_1^{(k)}, \ldots, \mathcal{E}_{\nu(k)}^{(k)}, \Delta_k$ of the state space E such that for all $1 \leq k \leq \mathfrak{M}$:*

(P1) $1 < \nu(k) < \nu(k-1)$, $k \geq 2$.
(P2) *For $k \geq 2$, $1 \leq i \leq \nu(k)$, $\mathcal{E}_i^{(k)} = \cup_{a \in I_{k,i}} \mathcal{E}_a^{(k-1)}$, where $I_{k,1}, \ldots, I_{k,\nu(k)}$ are disjoint subsets of $\{1, \ldots, \nu(k-1)\}$.*
(P3) *For $1 \leq i \leq \nu(k)$, $\mu_N(x) \approx \mu_N(\mathcal{E}_i^{(k)})$ for all states x in $\mathcal{E}_i^{(k)}$.*
(P4) *Let $\mathcal{E}^{(k)} = \cup_{i=1}^{\nu(k)} \mathcal{E}_i^{(k)}$. For all $1 \leq i \neq j \leq \nu(k)$, the following limits exist:*

$$\mathfrak{r}_k(i, j) := \lim_{N \to \infty} \theta_N(k) \, r_N^{\mathcal{E}^{(k)}}(\mathcal{E}_i^{(k)}, \mathcal{E}_j^{(k)}).$$

(P5) *Property (M1′) of metastability holds: For every $1 \leq i \leq \nu(k)$, every state x in $\mathcal{E}_i^{(k)}$ and every $\delta > 0$,*

$$\lim_{N \to \infty} \max_{y \in \mathcal{E}_i^{(k)}} \mathbf{P}_y^N\big[T_x > \delta\theta_N(k)\big] = 0.$$

(P6) *Property (M2) of metastability holds: Let $\Psi_k : \mathcal{E}^{(k)} \to \{1, \ldots, \nu(k)\}$ be given by*

$$\Psi_k(x) = \sum_{i=1}^{\nu(k)} i \, \mathbf{1}\{x \in \mathcal{E}_i^{(k)}\}.$$

Denote by $\{\eta_t^{N,k} : t \geq 0\}$ the trace of the process $\{\eta_t^N : t \geq 0\}$ on $\mathcal{E}^{(k)}$. For every $1 \leq i \leq \nu(k)$, $x \in \mathcal{E}_i^{(k)}$, under the measure $\mathbf{P}_x^N$, the blind speeded-up (non-Markovian) process $X_t^{N,k} = \Psi_k(\eta_{t\theta_N(k)}^{N,k})$ converges to the Markov process on $\{1, \ldots, \nu(k)\}$ starting from i and characterized by the rates $\mathfrak{r}_k(i, j)$.
(P7) *Property (M3′) of metastability holds: For every $t > 0$,*

$$\lim_{N \to \infty} \max_{x \in E} \mathbf{E}_x^N\left[\int_0^t \mathbf{1}\{\eta_{s\theta_N(k)}^N \in \Delta_k\} \, ds\right] = 0.$$

The sets $\mathcal{E}_i^{(k)}$, $1 \leq i \leq \nu(k)$, are called the metastable states at level k or, simply, k-metastates. Property (P2) asserts that as we pass from a description in the time scale $\theta_N(k-1)$ to a description in the longer time scale $\theta_N(k)$, the new metastates are larger and obtained as unions of $(k-1)$-metastates. Moreover, by property (P3), all states in any metastable set have measure of the same magnitude.

Condition (P5) asserts that, with a probability increasing to 1, any state in a metastable set is visited before the process leaves the metastable set. The process therefore thermalizes in the metastable state or, in other words, reaches a local equilibrium, before leaving the metastable state.

Condition (P7) states that on the time scale $\theta_N(k)$, the time spent outside the union of all metastates is negligible. To examine the behavior of the process in this time scale we may therefore restrict our attention to the trace process $\{\eta_t^{N,k} : t \geq 0\}$ speeded up by $\theta_N(k)$.

It follows from properties (P5) and (P6) that the speeded-up trace process $\{\eta_{t\theta_N(k)}^{N,k} : t \geq 0\}$ thermalizes in each metastable set $\mathcal{E}_i^{(k)}$ and then, at the end of an exponential time, jumps to another metastable set. By property (P4) the rate at which the process jumps from one metastable set to another is given by the asymptotic mean rate at which the speeded-up trace process jumps. Theorem 2.1 gives, therefore, a complete description of the evolution of the process on each time scale $\theta_N(k)$.

Remark 2.2. In order to apply this result to concrete models, we proceed as follows. Consider the Markov process on E obtained by suppressing all jumps $R_N(x, y)$ of magnitude smaller than 1: $R_N(x, y) \prec 1$. Note that this Markov process may not be irreducible. Denote by $\nu = \nu(1)$ the number of irreducible classes and by $\mathcal{E}_1, \ldots, \mathcal{E}_\nu$ the irreducible classes. These sets are the 1-metastates. Let

$$\theta_{N,i} = \frac{\mu_N(\mathcal{E}_i)}{\mathrm{cap}_N(\mathcal{E}_i, \breve{\mathcal{E}}_i)}, \quad 1 \leq i \leq \nu, \tag{2.7}$$

where $\mathrm{cap}_N(A, B)$ represents the capacity between A and B, defined in Section 4, and $\breve{\mathcal{E}}_i = \cup_{j \neq i} \mathcal{E}_j$. By Proposition 5.8 the sequences $(\theta_{N,i} : N \geq 1), 1 \leq i \leq \nu$, are comparable. Let $\theta_N = \theta_N(1) = \min\{\theta_{N,i} : 1 \leq i \leq \nu\}$. Since the sequences are comparable the following limits exist:

$$\lambda(i) = \lim_{N \to \infty} \frac{\theta_N}{\theta_{N,i}} \in [0, 1], \quad 1 \leq i \leq \nu.$$

By Lemma 4.3 and the first remark formulated at the end of Section 6, for every $1 \leq i \neq j \leq \nu$, the limits below also exist and do not depend on $x \in \mathcal{E}_i$:

$$p(i, j) = \lim_{N \to \infty} \mathbf{P}_x^N \big[T_{\mathcal{E}_j} = T_{\breve{\mathcal{E}}_i} \big]. \tag{2.8}$$

By (6.2), $\mathfrak{r}_1(i, j) = \lambda(i) p(i, j)$.

Hence, the characterization of the 1-metastates is very simple and the computation of $\theta_{N,i}$ (the depth of the valley $\mathcal{E}_i$, as we shall see) is feasible. This computation provides the slowest time scale $\theta_N(1)$ in which a metastable behavior is observed. To determine the exact asymptotic evolution in this time scale, we need to compute (2.8) which may be difficult or even impossible. In several cases, however, one may at least discern the pairs (i, j) for which $\mathfrak{r}_1(i, j)$ is strictly positive. This permits us to iterate the argument and gives an imprecise picture of the metastable behavior. In the time scale $\theta_N(1)$ the process remains in the 1-metastate $\mathcal{E}_i$ for a rate $\lambda(i)$ exponential time at the end of which it jumps to one of the remaining metastates such that $p(i, j) > 0$.

Consider the Markov process on $\{\mathcal{E}_1, \ldots, \mathcal{E}_\nu\}$ (instead of $\{1, \ldots, \nu\}$) with rates $\mathfrak{r}_1(i, j)$ and denote by $\nu(2)$ the number of its irreducible classes, and by $\mathcal{E}_1^{(2)}, \ldots, \mathcal{E}_{\nu(2)}^{(2)}$ the irreducible classes. Note that properties (P1) and (P2) are fulfilled and that we only need to know whether

$p(i, j)$ is strictly positive or not to determine the irreducible classes. Compute (2.7) and (2.8) for this new class of sets to obtain the second time scale $\theta_N(2)$ and the rates $\mathfrak{r}_2(i, j)$. Iterating this scheme we completely characterize the metastable behavior of the Markov process.

We conclude this section with some comments. In statistical mechanics models, the rates $R_N(x, y)$ are usually exponential and given by $e^{Nh(x,y)}$ for some function $h : E \times E \to \mathbb{R}$. Assumptions (2.1), (2.2) are trivially satisfied in this context.

In some models examined in statistical mechanics the time scales $\theta_N(k)$, $1 \le k < \mathfrak{M}$, correspond to the nucleation phase of the system, which may be very intricate even for simple dynamics due to the variety of valleys and the complexity of their geometries. In most cases, one only investigates the behavior in the largest time scale, $\theta_N(\mathfrak{M})$, where one observes either an exponential jump from a metastable to a stable state, or a Markovian evolution among competing metastable states.

We prove in (6.3) and (7.1) that the process never jumps from a metastable set to another metastable set which has probability of smaller order: $\mathfrak{r}_k(i, j) = 0$ if $\mu_N(\mathcal{E}_j^{(k)}) \prec \mu_N(\mathcal{E}_i^{(k)})$.

3. The Ising model at low temperature

To illustrate the methods presented in the first part of this article, we examine in this section the metastable behavior of the Ising model at low temperature following Neves and Schonmann [20].

We consider the two-dimensional nearest neighbor ferromagnetic Ising model on a finite torus $\Lambda_L = \mathbb{T}_L \times \mathbb{T}_L$, $L \ge 1$, where $\mathbb{T}_L = \{1, \ldots, L\}$ is the discrete one-dimensional torus with L points. The Hamiltonian is written as

$$\mathbb{H}(\sigma) = -\frac{1}{2} \sum_{\langle x, y \rangle} \sigma(x)\sigma(y) - \frac{h}{2} \sum_{x \in \Lambda_L} \sigma(x),$$

where $\sigma(x) \in \{-1, 1\}$, the first sum runs over the pairs of nearest neighbor sites of Λ_L, counting each pair only once, and the second is taken over Λ_L. We will always consider $h > 0$.

At inverse temperature $\beta > 0$, the Gibbs measure μ_β associated with the Hamiltonian $\mathbb{H}$ is given by

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

where Z_β is the normalizing partition function.

The Glauber dynamics on the state space $\Omega = \Omega_L = \{-1, 1\}^{\Lambda_L}$, also known as the Ising model, is the continuous time Markov process whose generator L_β acts on functions $f : \Omega \to \mathbb{R}$ as

$$(L_\beta f)(\sigma) = \sum_{x \in \Lambda_L} c(x, \sigma) [f(\sigma^x) - f(\sigma)],$$

where σ^x is the configuration obtained from σ by flipping the spin at x:

$$\sigma^x(y) = \begin{cases} \sigma(y) & \text{if } y \ne x, \\ -\sigma(x) & \text{if } y = x, \end{cases}$$

where the rates $c(x, \sigma)$ are given by

$$c(x, \sigma) = \exp\{-\beta[\mathbb{H}(\sigma^x) - \mathbb{H}(\sigma)]_+\},$$

and where a_+, $a \in \mathbb{R}$, stands for the positive part of a: $a_+ = \max\{a, 0\}$. The Markov process $\{\sigma_t^\beta : t \geq 0\}$ with generator L_β is reversible with respect to the Gibbs measures μ_β, $\beta > 0$, and ergodic. Denote by $R_\beta(\sigma, \sigma')$ the rate at which the process jumps from σ to σ', so $R_\beta(\sigma, \sigma')$ vanishes unless $\sigma' = \sigma^x$ for some $x \in \Lambda_L$, in which case $R_\beta(\sigma, \sigma^x) = c(x, \sigma)$.

In this model the process jumps from a state σ to the state σ^x at rate 1 if $\mu_\beta(\sigma) \leq \mu_\beta(\sigma^x)$. In particular, by the detailed balance condition, $\mu_\beta(\sigma) R_\beta(\sigma, \sigma^x) = \min\{\mu_\beta(\sigma), \mu_\beta(\sigma^x)\}$ so $G_\beta(\sigma, \sigma^x) = \min\{\mu_\beta(\sigma), \mu_\beta(\sigma^x)\}$.

We examine in this section the metastable behavior of the Markov process $\{\sigma_t^\beta : t \geq 0\}$ on Ω as the temperature vanishes. To avoid degenerate cases, following [20] we assume from now on that $0 < h < 2$, that $2/h \notin \mathbb{N}$ and that $L > (n_0 + 1)^2 + 1$, where $n_0 = [2/h]$ and $[r]$ stands for the integer part of r. We refer the reader to [19] for the degenerate case $2/h \in \mathbb{N}$.

Let I be an interval of the one-dimensional torus $\mathbb{T}_L$. The sets $I \times \mathbb{T}_L, \mathbb{T}_L \times I \subset \Lambda_L$ are called rings, while rectangles are subsets of the form $I \times J$, where I, J are non-empty proper intervals of $\mathbb{T}_L$.

To describe all metastable behaviors of the Ising model, we need to define the time scales at which they occur, the metastable sets associated with each time scale, and the asymptotic dynamics which specifies at which rate the process jumps from one metastable state to another. We start defining the $n_0 + 1$ time scales. For $1 \leq k \leq n_0 - 1$ let

$$\theta_\beta(k) = e^{k\beta h}, \qquad \theta_\beta(n_0) = e^{\beta(2-h)}, \qquad \theta_\beta(n_0 + 1) = e^{\beta c(h)},$$

where $c(h) = 4(n_0 + 1) - h[(n_0 + 1)n_0 + 1]$. Note that $\theta_\beta(1) \prec \cdots \prec \theta_\beta(n_0 + 1)$.

The presentation of the metastable sets requires some notation. Denote by $\Omega_o \subset \Omega$ the set of configurations whose total jump rate $\sum_{x \in \Lambda_L} R_\beta(\sigma, \sigma^x)$ vanishes as $\beta \uparrow \infty$. This is the set of configurations in which a negative spin has at most one positive neighbor and in which a positive spin has at most two negative neighbors. This set contains the configurations $+\mathbf{1}, -\mathbf{1}$, which are the configurations with all spins positive, negative, respectively, and configurations formed by positive rectangles and rings of length and width larger than 2 in a background of negative spins. In these latter configurations, to fulfill the prescribed conditions the positive rectangles and rings may not be at graph distance 2.

For a configuration σ in Ω_o, denote by $\ell(\sigma)$ the smallest length or width of the positive rectangles of σ. By convention, $\ell(-\mathbf{1}) = 0$, $\ell(+\mathbf{1}) = L$ and $\ell(\sigma) = L$ if σ contains no positive rectangles, but only positive rings. Let $N_r(\sigma)$ be the number of positive $\ell(\sigma) \times m$ rectangles of σ for some $m > \ell(\sigma)$, and let $N_s(\sigma)$ be the number of positive $\ell(\sigma) \times \ell(\sigma)$ squares of σ.

We may now introduce the metastable states $\Omega_{o,k}$ appearing in the time scale $\theta_\beta(k)$, $1 \leq k \leq n_0 + 1$. For $1 \leq k \leq n_0$, let

$$\Omega_{o,k} = \{\sigma \in \Omega_o : \ell(\sigma) > k\} \cup \{-\mathbf{1}\}, \qquad \Omega_{o,n_0+1} = \{+\mathbf{1}, -\mathbf{1}\}.$$

Note that $\Omega_o = \Omega_{o,1} \supset \cdots \supset \Omega_{o,n_0+1}$. The metastable states appearing in the time scale $\theta_\beta(k)$, $1 \leq k \leq n_0 + 1$, are all the elements of $\Omega_{o,k}$.

To depict how the process jumps from one metastable state to another in the different time scales, we need to introduce several sets. We use the terminology of graph theory to name some of them. Denote by $\mathbb{D}(\sigma)$ the set of direct successors of the configuration σ in $\Omega_o, \sigma \neq +\mathbf{1}, -\mathbf{1}$. If $\ell(\sigma) = 2$, $\mathbb{D}(\sigma)$ is the set of configurations obtained from σ by flipping all positive spins from one of the two sides of length 2 of a positive $2 \times m$ rectangle, $m > 2$, and of configurations obtained from σ by flipping all spins of a positive 2×2 square of σ. Clearly, $|\mathbb{D}(\sigma)| = 2N_r(\sigma) + N_s(\sigma)$. When $3 \leq \ell(\sigma) \leq n_0$, $\mathbb{D}(\sigma)$ is the set of configurations obtained

from σ by flipping all positive spins from one of the sides of length $\ell(\sigma)$ of a positive $\ell(\sigma) \times m$ rectangle, $m \geq \ell(\sigma)$. In this case, $|\mathbb{D}(\sigma)| = 2N_r(\sigma) + 4N_s(\sigma)$. For $\ell(\sigma) > n_0$, $\mathbb{D}(\sigma)$ is the set of configurations obtained by first flipping a negative spin from a site which has a neighbor site with a positive spin, and then flipping in any order all negative spins surrounded by two positive spins. Note that in this latter case two or more positive rectangles may be replaced by the smallest rectangle which contains them all. For this reason an exact description of the direct successors of a configuration in the case $\ell(\sigma) > n_0$ is more complicated.

For $\sigma \in \Omega_o$, $\sigma \neq \pm \mathbf{1}$, denote by $\mathbb{W}(\sigma)$ the set of saddle points of the configuration σ. For $2 \leq \ell(\sigma) \leq n_0$, $\mathbb{W}(\sigma)$ is the set of configurations obtained from σ by flipping $\ell(\sigma) - 1$ positive spins from a side of length $\ell(\sigma)$ of a positive $\ell(\sigma) \times m$ rectangle of σ, $m \geq \ell(\sigma)$. Note that $|\mathbb{W}(\sigma)| = 2\ell(\sigma)N_r(\sigma) + 4\ell(\sigma)N_s(\sigma)$ for $3 \leq \ell(\sigma) \leq n_0$ and $|\mathbb{W}(\sigma)| = 4N_r(\sigma) + 4N_s(\sigma)$ for $\ell(\sigma) = 2$. For $\ell(\sigma) > n_0$, $\mathbb{W}(\sigma)$ consists of the set of configurations obtained from σ by flipping a negative spin from a site which has one neighbor with a positive spin such that $|\mathbb{W}(\sigma)|$ is equal to the sum of the perimeters of the positive rectangles of σ added to $2L$ times the number of positive rings of σ.

For $\ell(\sigma) > n_0$, let $\mathbb{W}(\sigma, \sigma')$, $\sigma \in \Omega_o$, $\sigma' \in \mathbb{D}(\sigma)$, be the subset of $\mathbb{W}(\sigma)$ of all configurations which attain σ' by flipping in any order all negative spins surrounded by two positive spins, and let $\mathbb{W}_j(\sigma)$, $1 \leq j \leq 3$, be the configurations ξ of $\mathbb{W}(\sigma)$ with the following property. The site where ξ differs from σ has three neighbors with negative spins. Among these three neighbors, j sites have two neighbors with positive spins. The case $j = 3$ occurs when the configuration has two positive rectangles or rings at distance 3. Let $\mathbb{W}_j(\sigma, \sigma') = \mathbb{W}_j(\sigma) \cap \mathbb{W}(\sigma, \sigma')$.

Fix a configuration $\sigma \in \Omega_o$ and let $\Omega_\sigma = \Omega_o \setminus \{\sigma\}$. Recall that we denote by T_A the time of hitting of a set $A \subset \Omega$. We prove in Lemma 8.2 that $\mathbf{P}_\sigma^\beta[T_{\mathbb{D}(\sigma)} = T_{\Omega_\sigma}]$ converges, as $\beta \uparrow \infty$, to 1 and that the process reaches σ' by first visiting a configuration of $\mathbb{W}(\sigma)$.

Denote by $\mathbb{S}(\sigma)$ the set of successors of the configuration σ in Ω_o, $\sigma \neq +\mathbf{1}, -\mathbf{1}$. The difference between a successor and a direct successor is that the critical length $\ell(\sigma')$ of a successor σ' may not be smaller than that of the original configuration: $\ell(\sigma') \geq \ell(\sigma)$. If $\ell(\sigma) = 2$ or $\ell(\sigma) > n_0$, the set of successors coincides with the set of direct successors: $\mathbb{S}(\sigma) = \mathbb{D}(\sigma)$. However, if $3 \leq \ell(\sigma) \leq n_0$, $\mathbb{S}(\sigma)$ is the set of configurations obtained from σ by flipping all positive spins from one of the two sides of length $\ell(\sigma)$ of a positive $\ell(\sigma) \times m$ rectangle of σ, $m > \ell(\sigma)$, and of configurations obtained from σ by flipping all spins of a positive $\ell(\sigma) \times \ell(\sigma)$ square of σ.

The probability measure p introduced below describes how the process jumps from one metastable state to another in the appropriate time scales. For each configuration $\sigma \in \Omega_o$, define the probability measure $p(\sigma, \cdot)$ on Ω_o as follows. Let $p(\sigma, \sigma') = 0$ for $\sigma' \notin \mathbb{S}(\sigma)$. For $\sigma' \in \mathbb{S}(\sigma)$ and $\ell(\sigma) = 2 \leq n_0$, let

$$p(\sigma, \sigma') = \begin{cases} (8/3)[2N_r + (8/3)N_s]^{-1} & \text{for } \sigma' \in \mathbb{S}_s(\sigma), \\ [2N_r + (8/3)N_s]^{-1} & \text{otherwise,} \end{cases} \tag{3.1}$$

where $\mathbb{S}_s(\sigma) \subset \mathbb{S}(\sigma)$ is the set of configurations obtained from σ by flipping all spins in a positive 2×2 square of σ. For $\sigma' \in \mathbb{S}(\sigma)$ and $3 \leq \ell(\sigma) \leq n_0$, let

$$p(\sigma, \sigma') = \begin{cases} 4[2N_r + 4N_s]^{-1} & \text{for } \sigma' \in \mathbb{S}_s(\sigma), \\ [2N_r + 4N_s]^{-1} & \text{otherwise,} \end{cases} \tag{3.2}$$

where $\mathbb{S}_s(\sigma) \subset \mathbb{S}(\sigma)$ is the set of configurations obtained from σ by flipping all spins in a positive $\ell(\sigma) \times \ell(\sigma)$ square of σ. Finally, for $\sigma' \in \mathbb{S}(\sigma)$ and $\ell(\sigma) > n_0$, let

$$p(\sigma, \sigma') = \frac{\sum_{j=1}^{3} \frac{j}{j+1} |\mathbb{W}_j(\sigma, \sigma')|}{\sum_{j=1}^{3} \frac{j}{j+1} |\mathbb{W}_j(\sigma)|}. \tag{3.3}$$

It remains to describe the rates at which the process leaves a metastable state in the different time scales. Let $\theta : \Omega_o \setminus \{-\mathbf{1}, +\mathbf{1}\} \to \mathbb{R}_+$ be given by

$$\theta(\sigma) = \begin{cases} (2/3)N_s(\sigma) + 2N_r(\sigma) & \text{if } \ell = 2 \le n_0, \\ \dfrac{2\ell - 1}{3\ell} |\mathbb{W}(\sigma)| & \text{if } 3 \le \ell \le n_0, \\ (1/2)|\mathbb{W}_1(\sigma)| + (2/3)|\mathbb{W}_2(\sigma)| + (3/4)|\mathbb{W}_3(\sigma)| & \text{if } \ell > n_0. \end{cases} \tag{3.4}$$

We are now in a position to state the first main result of this section. Fix $1 \le k \le n_0 + 1$ and denote by $\sigma_t^{\beta,k}$ the trace of the process σ_t^{β} on $\Omega_{o,k}$. Recall that $\sigma_t^{\beta,k}$ is a Markov process on $\Omega_{o,k}$.

Theorem 3.1. *Fix $1 \le k \le n_0$. As $\beta \uparrow \infty$, the Markov process $\sigma_{t\theta_\beta(k)}^{\beta,k}$ converges to the Markov process on $\Omega_{o,k}$ with jump rates $\mathfrak{r}$ given by*

$$\mathfrak{r}(\sigma, \sigma') = \begin{cases} \theta(\sigma)p(\sigma, \sigma') & \text{if } \sigma \in \Omega_{o,k} \setminus \Omega_{o,k+1}, \\ 0 & \text{if } \sigma \in \Omega_{o,k+1}. \end{cases}$$

Moreover, the time spent outside $\Omega_{o,k}$ by the process $\sigma_{t\theta_\beta(k)}^{\beta,k}$ is negligible: for all $t > 0$ and $\sigma \in \Omega$,

$$\lim_{\beta \to \infty} \mathbf{E}_\sigma^\beta \left[\int_0^t \mathbf{1}\{\sigma_{s\theta_\beta(k)}^{\beta,k} \notin \Omega_{o,k}\} \, ds \right] = 0.$$

Fix a configuration $\sigma \in \Omega_{o,k}$, $1 \le k \le n_0 - 1$, and consider asymptotic behavior, as the temperature vanishes, of the trace process $\sigma_t^{\beta,k}$ in the time scale $\theta_\beta(k)$ starting from σ. Theorem 3.1 states that if $\ell(\sigma) > k + 1$, the configuration σ is an absorbing point for the asymptotic dynamics, while if $\ell(\sigma) = k + 1$, the asymptotic dynamics visits a sequence of configurations where each element of the sequence differs from the previous one either by flipping all positive spins of one of the two sides of length $k + 1$ of a $(k + 1) \times m$ positive rectangle, $m > k + 1$, or by flipping all spins of a $(k + 1) \times (k + 1)$ positive square. After a finite number of jumps, the process reaches a configuration whose positive rectangles have all sides larger than $k + 1$ and stays there forever.

For a configuration $\sigma \in \Omega_{o,n_0}$, Theorem 3.1 states that in the time scale $\theta_\beta(n_0)$ the trace process σ_t^{β,n_0} sees its positive rectangles and rings increase gradually until the configurations $+\mathbf{1}$ is reached.

This result describes therefore the behavior of the Ising model in the intermediate scales where first the small positive droplets are removed and then the large positive droplets increase to eventually occupy all space. To complete the picture of the metastable behavior of the model it remains to specify how the process jumps from the configuration $-\mathbf{1}$ to the configuration $+\mathbf{1}$.

Denote by $\mathbb{W}(-\mathbf{1})$ the set of configurations which have a positive $(n_0 + 1) \times n_0$ rectangle and an extra positive spin which has a positive neighbor sitting on one of the sides of length $n_0 + 1$ of the positive rectangle, all others spins being negative. All configurations of $\mathbb{W}(-\mathbf{1})$ have the same measure. Denote by $\mathbb{W}_1(-\mathbf{1})$ the configurations of $\mathbb{W}(-\mathbf{1})$ whose extra positive spin is next to the corner of the positive rectangle and by $\mathbb{W}_2(-\mathbf{1})$ the other configurations of $\mathbb{W}_1(-\mathbf{1})$. Let

$$\theta(-\mathbf{1}) = (1/2)|\mathbb{W}_1(-\mathbf{1})| + (2/3)|\mathbb{W}_2(-\mathbf{1})|.$$

Theorem 3.2. *As $\beta \uparrow \infty$, the Markov process $\sigma_{t\theta_\beta(n_0+1)}^{\beta,n_0+1}$ converges to the Markov process on $\{-\mathbf{1}, +\mathbf{1}\}$ in which $+\mathbf{1}$ is an absorbing state and which jumps from $-\mathbf{1}$ to $+\mathbf{1}$ at rate $\theta(-\mathbf{1})$. Moreover, the time spent outside $\{-\mathbf{1}, +\mathbf{1}\}$ by the process $\sigma_{t\theta_\beta(n_0+1)}^{\beta,n_0+1}$ is negligible: for all $t > 0$ and $\sigma \in \Omega$,*

$$\lim_{\beta \to \infty} \mathbf{E}_\sigma^\beta \left[\int_0^t \mathbf{1}\{\sigma_{s\theta_\beta(n_0+1)}^{\beta,n_0+1} \neq \pm \mathbf{1}\} \, \mathrm{d}s \right] = 0.$$

4. Capacities and hitting times

Denote by D_N the Dirichlet form associated with the generator of the Markov process introduced in Section 2:

$$D_N(f) = \sum_{\{x,y\} \subset E} \mu_N(x) R_N(x, y) \{f(y) - f(x)\}^2, \quad f : E \to \mathbb{R},$$

where in the sum on the right hand side each bond $\{x, y\}$ is counted only once. Let $\mathrm{cap}_N(A, B)$, $A, B \subset E$, $A \cap B = \varnothing$, be the capacity between A and B:

$$\mathrm{cap}_N(A, B) = \inf_f D_N(f), \tag{4.1}$$

where the infimum is taken over all functions $f : E \to \mathbb{R}$ such that $f(x) = 1$ for all $x \in A$, and $f(x) = 0$ for all $x \in B$.

We may compute the order of magnitude of the capacity between two disjoint subsets of E. A self-avoiding path γ from A to B, $A, B \subset E$, $A \cap B = \varnothing$, is a sequence of sites $(x_0, x_1, \ldots, x_n)$ such that $x_0 \in A$, $x_n \in B$, $x_i \neq x_j$, $i \neq j$, $R_N(x_i, x_{i+1}) > 0$, $0 \leq i < n$. Denote by $\Gamma_{A,B}$ the set of self-avoiding paths from A to B and let

$$G_N(A, B) := \max_{\gamma \in \Gamma_{A,B}} G_N(\gamma), \qquad G_N(\gamma) := \min_{0 \leq i < n} G_N(x_i, x_{i+1}).$$

Note that there might be more than one optimal path and that $G_N(\{x\}, \{y\}) \geq G_N(x, y)$, with possibly a strict inequality.

We shall say that a bond (x_p, x_{p+1}) of a path $\gamma = (x_0, x_1, \ldots, x_n)$ is *critical* if

$$G_N(x_p, x_{p+1}) = \min_{0 \leq i < n} G_N(x_i, x_{i+1}) = G_N(\gamma).$$

Note that for all disjoint sets A, B, C,

$$G_N(A, B \cup C) = \max\{G_N(A, B), G_N(A, C)\}. \tag{4.2}$$

Indeed, the left hand side is greater than or equal to the right hand side because $G_N(A, D) \leq G_N(A, D')$ if $D \subset D'$. On the other hand, there exists a self-avoiding path $\gamma = (x_0, \ldots, x_n)$

from A to $B \cup C$ such that $G_N(A, B \cup C) = G_N(\gamma)$. We may assume without loss of generality that x_n belongs to B. Hence, γ is a self-avoiding path from A to B and $G_N(\gamma) \leq G_N(A, B)$, which proves (4.2).

Lemma 4.1. *Fix $A, B \subset E$ such that $A \cap B = \varnothing$. The capacity $\mathrm{cap}_N(A, B)$ is of the same magnitude as $G_N(A, B)$. There exists a positive and finite constant C_1, depending only on E and on the limiting rates $r(x, y)$, such that*

$$C_1^{-1} \leq \frac{\mathrm{cap}_N(A, B)}{G_N(A, B)} \leq C_1$$

for all N sufficiently large.

Proof. Fix two subsets A, B of E such that $A \cap B = \varnothing$. We shall obtain an upper and a lower bound for $\mathrm{cap}_N(A, B)$. We start with a lower bound. Fix a self-avoiding path $\gamma = (x_0, x_1, \ldots, x_n)$ in $\Gamma_{A,B}$ such that $G_N(A, B) = \min_{0 \leq i < n} G_N(x_i, x_{i+1})$. This path always exists because the number of paths is finite. For any function $f : E \to \mathbb{R}$,

$$D_N(f) \geq \sum_{i=0}^{n-1} G_N(x_i, x_{i+1})[f(x_{i+1}) - f(x_i)]^2.$$

Therefore, minimizing over all $f : E \to \mathbb{R}$ such that $f(x) = 1$, $x \in A$, $f(y) = 0$, $y \in B$, we get that $\mathrm{cap}_N(A, B)$ is bounded below by

$$\inf_f \sum_{i=0}^{n-1} G_N(x_i, x_{i+1})[f(x_{i+1}) - f(x_i)]^2,$$

where the infimum is taken over all functions $f : \{x_0, \ldots, x_n\} \to \mathbb{R}$ such that $f(x_0) = 1$, $f(x_n) = 0$. A simple computation shows that this expression is equal to

$$\left\{ \sum_{i=0}^{n-1} \frac{1}{G_N(x_i, x_{i+1})} \right\}^{-1},$$

which is bounded below, for N large, by $C_1 \min_{0 \leq i < n} G_N(x_i, x_{i+1})$ for some positive constant C_1 depending only on E and the asymptotic rates $r(x, y)$. By the definition of the path $\gamma = (x_0, x_1, \ldots, x_n)$, $\min_{0 \leq i < n} G_N(x_i, x_{i+1}) = G_N(A, B)$, which proves the lower bound for the capacity.

We now turn to the upper bound. Denote by $\mathfrak{B}_N$ the set of bonds (x, y) such that $G_N(x, y) > G_N(A, B)$. The state space E can be written as the disjoint union of maximal connected components. More precisely, there exist disjoint subsets $A_1, \ldots, A_m$ of E, possibly singletons, fulfilling the next three conditions:

- $E = \cup_{1 \leq j \leq m} A_j$;
- for any $x, y \in A_j$, there exists a path $\gamma = (x = x_0, x_1, \ldots, x_p = y)$ such that $G_N(x_i, x_{i+1}) > G_N(A, B)$ for all $0 \leq i < p$;
- for any $x \in A_j$, $y \in A_k$, $j \neq k$, there does not exist a path $\gamma = (x = x_0, x_1, \ldots, x_p = y)$ such that $G_N(x_i, x_{i+1}) > G_N(A, B)$ for all $0 \leq i < p$.

Note that if $A_j \cap A \neq \varnothing$ then $A_j \cap B = \varnothing$. Otherwise, there would be a self-avoiding path $(x_0, \ldots, x_n)$ from A to B such that $G_N(x_i, x_{i+1}) > G_N(A, B)$ for all $0 \leq i < n$, in contradiction with the definition of $G_N(A, B)$.

Consider a self-avoiding path $\gamma = (x_0, x_1, \ldots, x_n)$ in $\Gamma_{A,B}$ such that $G_N(A, B) = \min_{0 \leq i < n} G_N(x_i, x_{i+1})$. The path γ may have bonds (x_i, x_{i+1}) in $\mathfrak{B}_N$. We claim, however, that there exists a bond (x_q, x_{q+1}) in γ such that $G_N(x_q, x_{q+1}) = G_N(A, B)$ and such that there is no maximal connected component A_j of E such that

$$A_j \cap [A \cup \{x_0, \ldots, x_q\}] \neq \varnothing \quad \text{and} \quad A_j \cap [B \cup \{x_{q+1}, \ldots, x_n\}] \neq \varnothing. \tag{4.3}$$

To prove this claim, let $L \geq 1$ be the number of critical bonds in γ and fix a critical bond (x_p, x_{p+1}) for which (4.3) does not hold. There exists therefore a maximal connected component A_j of E such that $A_j \cap [A \cup \{x_0, \ldots, x_p\}] \neq \varnothing$ and $A_j \cap [B \cup \{x_{p+1}, \ldots, x_n\}] \neq \varnothing$. By overlapping the bond (x_p, x_{p+1}) with a path in A_j, we construct a new self-avoiding path $\gamma' = (x'_0, \ldots, x'_{n'})$ from A to B with possibly different initial or final point which avoids the bond (x_p, x_{p+1}).

Since all bonds which belong to γ' and not to γ are in $\mathfrak{B}_N$ and since $G_N(x_i, x_{i+1}) \geq G_N(x_p, x_{p+1}) = G_N(A, B)$, $0 \leq i < n$, $G_N(x'_i, x'_{i+1}) \geq G_N(A, B)$ for all $0 \leq i < n'$. On the other hand, since γ' is a self-avoiding path from A to B, $\min_{0 \leq i < n'} G_N(x'_i, x'_{i+1}) \leq G_N(A, B)$. Hence, $\min_{0 \leq i < n'} G_N(x'_i, x'_{i+1}) = G_N(A, B)$.

On the other hand, since all bonds which belong to γ' and not to γ are in $\mathfrak{B}_N$ and since (x_p, x_{p+1}) does not belong γ', the number of critical bonds of γ' is at most $L - 1$. It might be smaller than $L - 1$ if the set A_j overlaps more than one critical bond of γ.

If the new path γ' fulfills condition (4.3), the claim is proved. If it does not, we apply the algorithm again. Since the algorithm reduces the number of critical bonds by at least 1, after a finite number of iterations we obtain a path satisfying (4.3) as claimed.

We now define a function f equal to 1 on the set A and equal to 0 on the set B, and we show that the Dirichlet form of f is bounded by $C_1 G_N(A, B)$ for some finite constant C_1 which depends only on E. Let (x_p, x_{p+1}) be a critical bond of a path $\gamma = (x_0, \ldots, x_n)$ satisfying condition (4.3). Define $f : E \to \mathbb{R}$ as follows. Let $f(x) = 1$ for $x \in A$, $f(y) = 0, y \in B$. Define f on γ as $f(x_i) = 1$, $0 \leq i \leq p$, $f(x_j) = 0$, $p + 1 \leq i \leq n$. On each maximal connected component A_j which intersects $A \cup \{x_0, \ldots, x_p\}$, set $f = 1$. Similarly, on each maximal connected component A_k which intersects $\{x_{p+1}, \ldots, x_n\} \cup B$ set $f = 0$. Property (4.3) ensures that this can be done. For the remaining sites we define f to be a fixed arbitrary constant ω. Note that with this definition, f is constant on each maximal connected component A_k.

It remains to examine the Dirichlet form of f. There are three kinds of non-vanishing terms in this Dirichlet form. The first one is $G_N(x_p, x_{p+1}) = G_N(A, B)$. The second and third kinds are expressions of the form $G_N(x, y)(1 - \omega)^2, G_N(x, y)\omega^2$, where (x, y) does not belong to $\mathfrak{B}_N$. In particular, the contribution to the Dirichlet form of f of these expressions is bounded by $C_1 G_N(A, B)\{\omega^2 + (1 - \omega)^2\}$ for some finite constant which depends only on E. This proves that $D_N(f) \leq C_1 G_N(A, B)$. Since f is equal to 1 on the set A and is equal to 0 on the set B, $\mathrm{cap}_N(A, B) \leq C_1 G_N(A, B)$, which proves the lemma. $\quad\square$

This lemma presents a typical estimation of asymptotic capacities. We first obtain a lower bound of the Dirichlet form, uniform over all functions f, by disregarding some bonds. Then, we prove an upper bound for a specific candidate, believed to be close to the optimal function in view of the proof of the lower bound. This time, however, no bond can be neglected in the Dirichlet form.

Of course, the function f proposed in the proof of the previous lemma gives only the correct magnitude of the capacity $\mathrm{cap}_N(A, B)$ and not its exact asymptotic value. The computation of the exact asymptotic value requires detailed information on the jump rates and has to be done model by model.

We may prove, however, that under certain assumptions the capacity between two sets conveniently rescaled converges. Fix two disjoint subsets of $E : A, B \subset E, A \cap B = \varnothing$. By definition, $G_N(A, B) = G_N(x, y)$ for some $x, y \in E$. By (2.5), $G_N(x, y) = g_N(x, y) G_N(\mathfrak{j})$ for some $0 \leq j \leq \mathfrak{j}$. Let $\mathfrak{g}_N(A, B) = G_N(\mathfrak{j}) \leq 1$ so that $G_N(A, B)/\mathfrak{g}_N(A, B)$ converges, as $N \uparrow \infty$, to some number in $(0, \infty)$.

Lemma 4.2. *Fix two disjoint subsets of $E : A, B \subset E, A \cap B = \varnothing$. Let $f_N : E \to [0, 1]$ be the function $f_N(x) = \mathbf{P}_x^N[T_A < T_B]$. Assume that f_N converges pointwise to some function f. Denote by $\mathfrak{B}(A, B)$ the set of pairs $\{x, y\}$ such that $G_N(x, y) \approx \mathfrak{g}_N(A, B)$. Then, $f(y) = f(x)$ if $G_N(x, y) \succ \mathfrak{g}_N(A, B)$ and*

$$\lim_{N \to \infty} \frac{\text{cap}_N(A, B)}{\mathfrak{g}_N(A, B)} = \sum_{\{x,y\} \in \mathfrak{B}(A,B)} g(x, y)[f(y) - f(x)]^2 \in (0, \infty),$$

where $g(x, y)$ has been introduced in (2.5). In particular, $\text{cap}_N(A, B)/G_N(A, B)$ converges, as $N \uparrow \infty$, to a finite and strictly positive limit:

$$\lim_{N \to \infty} \frac{\text{cap}_N(A, B)}{G_N(A, B)} =: \varXi(A, B) \in (0, \infty).$$

Proof. Fix two disjoint subsets of $E : A, B \subset E, A \cap B = \varnothing$ and let $f_N : E \to [0, 1]$ be the function $f_N(x) = \mathbf{P}_x^N[T_A < T_B]$. It is well known that

$$\text{cap}_N(A, B) = D_N(f_N) = \sum_{\{x,y\} \subset E} G_N(x, y)[f_N(y) - f_N(x)]^2.$$

We first show that $f(y) = f(x)$ if $G_N(x, y) \succ \mathfrak{g}_N(A, B)$. Indeed, fix such a pair and note that

$$G_N(x, y)[f_N(y) - f_N(x)]^2 \leq \text{cap}_N(A, B).$$

By Lemma 4.1, the right hand side is bounded above by $C_1 \mathfrak{g}_N(A, B)$ for some finite constant C_1 independent of N. Since f_N converges to f pointwise and since $G_N(x, y) \succ \mathfrak{g}_N(A, B)$, $f(y) = f(x)$, proving the claim.

Let $\mathfrak{B} = \mathfrak{B}(A, B)$. To prove a lower bound for the capacity, note that

$$\text{cap}_N(A, B) \geq \sum_{\{x,y\} \in \mathfrak{B}} G_N(x, y)[f_N(y) - f_N(x)]^2.$$

In view of (2.5), as $N \uparrow \infty$ the right hand side divided by $\mathfrak{g}_N(A, B)$ converges to

$$\sum_{\{x,y\} \in \mathfrak{B}} g(x, y)[f(y) - f(x)]^2.$$

To prove the upper bound, recall the variational formula (4.1) for the capacity to write

$$\text{cap}_N(A, B) \leq \sum_{\{x,y\} \subset E} G_N(x, y)[f(y) - f(x)]^2.$$

Since $f(y) = f(x)$ if $G_N(x, y) \succ \mathfrak{g}_N(A, B)$, and since f is absolutely bounded by 1, we may restrict the sum to the pairs (x, y) in $\mathfrak{B}(A, B)$. Hence,

$$\limsup_{N \to \infty} \frac{\text{cap}_N(A, B)}{\mathfrak{g}_N(A, B)} \leq \sum_{\{x,y\} \in \mathfrak{B}} g(x, y)[f(y) - f(x)]^2,$$

which proves the second assertion of the lemma. Moreover, by Lemma 4.1 and since $G_N(A, B) \approx \mathfrak{g}_N(A, B)$, the limit belongs to $(0, \infty)$.

The last assertion of the lemma follows from the fact, mentioned just before the statement of the lemma, that $G_N(A, B)/\mathfrak{g}_N(A, B)$ converges, as $N \uparrow \infty$, to a finite and strictly positive limit. $\square$

This result shows that the sequence of capacities are comparable if the sequence of hitting functions $f_{A,B}^N(x) = \mathbf{P}_x^N[T_A < T_B]$ converge. This remark highlights the interest of the next result. Recall that we denote by $R_N^F(x, y), x, y \in F$, the jump rates of the trace process $\{\eta_t^F : t \geq 0\}, F \subset E$.

Lemma 4.3. *For every subset F of E, the sequences $(R_N^F(x, y) : N \geq 1)$, $x \neq y \in F$, are comparable. Moreover, for all subsets A, B of E, $A \cap B = \varnothing$, and for every $x \in E$, the following limits exist:*

$$f_{A,B}(x) := \lim_{N \to \infty} \mathbf{P}_x^N[T_A < T_B].$$

Proof. It follows from the displayed formula presented just after Corollary 6.2 in [2] that

$$R_N^F(x, y) = \frac{\sum\limits_{z \in E} R_N(x, y) R_N(w, z) + R_N(x, w) R_N(w, y)}{\sum\limits_{z \in E} R_N(w, z)}$$

if $F = E \setminus \{w\}$. Iterating this formula, we may show that for every proper subset F of E, $R_N^F(x, y)$ may be expressed as a ratio of sums of products of the rates $R_N(\cdot, \cdot)$. The sum in the numerator contains only products with the same number of terms and the same thing happens in the denominator. In particular, by assumption (2.2), the sequences $\{R_N^F(x, y) : N \geq 1\}, x \neq y \in F$, are comparable. This proves the first assertion of the lemma.

If we denote by $p_N^F(x, y)$ the jump probabilities associated with the rates $R_N^F(x, y)$,

$$p_N^F(x, y) = \frac{R_N^F(x, y)}{\sum\limits_{z \in F} R_N^F(x, z)}, \quad x \neq y \in F,$$

$p_N^F(x, y)$ converges to some $p^F(x, y)$ as $N \uparrow \infty$.

Denote by $\mathbf{P}_x^{N,F}, x \in F$, the probability on the path space $D(\mathbb{R}_+, F)$ induced by the trace process $\{\eta_t^F : t \geq 0\}$ starting from x. Clearly, $\mathbf{P}_x^N[T_A < T_B] = \mathbf{P}_x^{N,F}[T_A < T_B]$, for $F = \{x\} \cup A \cup B$. If x does not belong to $A \cup B$, the last probability is equal to $\sum_{y \in A} p_N^F(x, y)$ and we proved that this expression converges as $N \uparrow \infty$. $\square$

Corollary 4.4. *For every subset F of E and all subsets A, B of F, $A \cap B = \varnothing$, the ratio of mean rates*

$$\frac{r_N^F(A, B)}{r_N^F(A, F \setminus A)} = \frac{\sum\limits_{x \in A} \sum\limits_{y \in B} \mu^N(x) R_N^F(x, y)}{\sum\limits_{x \in A} \sum\limits_{z \in F \setminus A} \mu^N(x) R_N^F(x, z)}$$

converges to some number $p_F(A, B) \in [0, 1]$ as $N \uparrow \infty$.

Proof. It follows from the explicit formula for the rates R_N^F, derived in the proof of the previous lemma, from Eq. (2.3) and from assumption (2.2) that the sequences $(\mu^N(x) R_N^F(x, y) : N \geq 1)$, $x \neq y \in F$, are comparable. The result is a simple consequence of this observation. $\quad\square$

5. The shallowest valleys

Recall the definition of a valley with an attractor introduced in [2]. To avoid long sentences, in this article we call a valley with an attractor simply a valley. We describe in this section the shallowest valleys and we show that their depths are comparable.

We shall say that there exists an *open path* from x to y if there exists a sequence $x = x_0, x_1, \ldots, x_n = y$ such that $R_N(x_i, x_{i+1}) \approx 1$, $0 \leq i < n$. Two sites $x \neq y$ are said to be equivalent, $x \sim y$, if there exist an open path from x to y and an open path from y to x. If we also declare any site to be equivalent to itself, $\sim$ is an equivalence relation. We denote by $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_\alpha$ the equivalence classes.

Some equivalence classes are connected to other equivalence classes by open paths. On drawing an arrow from a set $\mathcal{C}_i$ to a set $\mathcal{C}_j$ if there exist $x \in \mathcal{C}_i$, $y \in \mathcal{C}_j$ such that $R_N(x, y) \approx 1$, the set $\{\mathcal{C}_1, \ldots, \mathcal{C}_\alpha\}$ becomes an oriented graph with no directed loops. We denote by $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_\nu$ the leaves of this graph, in the terminology of graph theory, the equivalence classes with no successors. Denote by Δ the union of the remaining sets such that $\{\mathcal{E}_1, \ldots, \mathcal{E}_\nu, \Delta\}$ forms a partition of E:

$$E = \mathcal{E} \cup \Delta, \quad \mathcal{E} = \mathcal{E}_1 \cup \cdots \cup \mathcal{E}_\nu. \tag{5.1}$$

For $1 \leq i \leq \nu$, let $\breve{\mathcal{E}}_i$ be the union of all leaves except $\mathcal{E}_i$:

$$\breve{\mathcal{E}}_i = \bigcup_{j \neq i} \mathcal{E}_j.$$

By construction, all sites in an equivalence class $\mathcal{C}_j$ have probability of the same magnitude: there exists a finite, positive constant C_0 such that for all $1 \leq j \leq \alpha$,

$$C_0^{-1} \leq \frac{\mu_N(x)}{\mu_N(y)} \leq C_0, \quad x, y \in \mathcal{C}_j. \tag{5.2}$$

We may also estimate the capacity between two states in a leaf $\mathcal{E}_i$.

Lemma 5.1. *Fix $1 \leq i \leq \nu$. There exists a finite constant C_1, which depends only on E, such that for any $x \neq y$ in $\mathcal{E}_i$,*

$$C_1^{-1} \leq \frac{\mathrm{cap}_N(\{x\}, \{y\})}{\mu_N(\mathcal{E}_i)} \leq C_1.$$

Proof. Fix $1 \leq i \leq \nu$ and $x \neq y$ in $\mathcal{E}_i$. Consider a function $f : E \to \mathbb{R}$ such that $f(x) = 1$, $f(y) = 0$ and fix a self-avoiding open path $\gamma = (x = x_0, \ldots, x_n = y)$ from x to y. By the Schwarz inequality,

$$1 = [f(y) - f(x)]^2 = \left\{ \sum_{i=0}^{n-1} \{f(x_{i+1}) - f(x_i)\} \right\}^2$$

$$\leq \sum_{i=0}^{n-1} \mu_N(x_i) R_N(x_i, x_{i+1}) \{f(x_{i+1}) - f(x_i)\}^2 \sum_{i=0}^{n-1} \frac{1}{\mu_N(x_i) R_N(x_i, x_{i+1})}.$$

Therefore, $D_N(f)$ is bounded below by

$$\sum_{i=0}^{n-1} \mu_N(x_i) R_N(x_i, x_{i+1}) \{f(x_{i+1}) - f(x_i)\}^2 \geq \left\{ \sum_{i=0}^{n-1} \frac{1}{\mu_N(x_i) R_N(x_i, x_{i+1})} \right\}^{-1}.$$

Since γ is an open path, $R_N(x_i, x_{i+1})$ is of order 1. Hence, by (5.2), there exists a constant C_1 which depends only on E such that for any function $f : E \to \mathbb{R}$ such that $f(x) = 1$, $f(y) = 0$, $D_N(f) \geq C_1^{-1} \mu_N(\mathcal{E}_i)$. This proves that $\text{cap}_N(\{x\}, \{y\}) \geq C_1^{-1} \mu_N(\mathcal{E}_i)$.

To prove the reverse inequality, consider the function $f_* : E \to \mathbb{R}$ which is equal to 1 at x and is 0 elsewhere. Clearly,

$$D_N(f_*) = \mu_N(x) \sum_{z \neq x} R_N(x, z).$$

By hypothesis, $R_N(x, z) \leq C_0$, so $D_N(f_*) \leq C_0' \mu_N(\mathcal{E}_i)$, proving the lemma. □

Recall Theorem 2.6 of [2] which presents sufficient conditions for a triple to be a valley in the context of reversible Markov processes.

Fix a leaf $\mathcal{E}_i$, $1 \leq i \leq \nu$, and a site x in $\mathcal{E}_i$. Denote by $\mathcal{E}_i$ the set $\mathcal{E}_i$ as well as the constant sequence of sets $(\mathcal{E}_i, \mathcal{E}_i, \ldots)$ and by x not only the site x but also the constant sequence equal to x. This convention is used from now on without further notice. Denote by $\mathcal{B}_i$ the set of sites in Δ of measure of lower magnitude than $\mathcal{E}_i$: $\mathcal{B}_i = \{y \in \Delta : \mu_N(y) \prec \mu_N(\mathcal{E}_i)\}$. Note that $\mathcal{B}_i$ is the union of some equivalence classes.

Lemma 5.2. *Fix $1 \leq i \leq \nu$ and x in $\mathcal{E}_i$. The triple $(\mathcal{E}_i, \mathcal{E}_i \cup \mathcal{B}_i, x)$ is a valley of depth $\theta_{N,i} :=$ $\mu_N(\mathcal{E}_i)/\text{cap}_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c)$.*

Proof. By [2, Theorem 2.6], to show that $(\mathcal{E}_i, \mathcal{B}_i \cup \mathcal{E}_i, x)$ is a valley of depth $\mu_N(\mathcal{E}_i)/\text{cap}_N$ $(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c)$ we need to check that $\mu_N(\mathcal{B}_i)/\mu_N(\mathcal{E}_i)$ vanishes as $N \uparrow \infty$ and that

$$\lim_{N \to \infty} \frac{\text{cap}_N(\mathcal{E}_i, [\mathcal{B}_i \cup \mathcal{E}_i]^c)}{\text{cap}_N(x)} = 0, \tag{5.3}$$

where $\text{cap}_N(x) = \min_{y \in \mathcal{E}_i} \text{cap}_N(\{x\}, \{y\})$. The first condition follows from the definition of the set $\mathcal{B}_i$. The second one is simple to check. Fix a positive function $f : E \to \mathbb{R}$ bounded by 1 and constant in $\mathcal{E}_i$ and in $[\mathcal{B}_i \cup \mathcal{E}_i]^c$. In the expression for the Dirichlet form $D_N(f) = \sum_{y,z} G_N(y, z)[f(z) - f(y)]^2$, there are two kinds of non-vanishing terms. Either y belongs to $\mathcal{E}_i$ and we may estimate $G_N(y, z)[f(z) - f(y)]^2$ by $\mu_N(\mathcal{E}_i) \max_{y \in \mathcal{E}_i, z \notin \mathcal{E}_i} R_N(y, z)$, or y does not belong to $\mathcal{E}_i$ and we may estimate $G_N(y, z)[f(z) - f(y)]^2$ by $\mu_N(\mathcal{B}_i)$ because $R_N(y, z) \leq C_0$. Both expressions are of an order much smaller than that of $\mu_N(\mathcal{E}_i)$ because $\mathcal{E}_i$ has no successors. Therefore, (5.3) follows from Lemma 5.1, proving that $(\mathcal{E}_i, \mathcal{B}_i \cup \mathcal{E}_i, x)$ is a valley. □

The next lemma shows that a leaf is attained from any site in a time scale of magnitude 1. Recall that T_A, $A \subset E$, stands for the time of hitting of A.

Lemma 5.3. *There exists a finite constant C_0, independent of N, such that*

$$\max_{x \in \Delta} \mathbf{E}_x^N [T_{\mathcal{E}}] \leq C_0.$$

Proof. Denote by $\{\tau_j : j \geq 0\}$ the jump times of the Markov process $\{\eta_t^N : t \geq 0\}$:

$$\tau_0 = 0, \qquad \tau_{j+1} = \inf\{t > \tau_j : \eta_t^N \neq \eta_{\tau_j}^N\}, \quad j \geq 0.$$

Denote by $\{Y_k^N : k \geq 0\}$ the jump chain associated with the Markov process $\{\eta_t^N : t \geq 0\}$, i.e., the discrete time Markov chain formed by the successive sites visited by η_t^N:

$$Y_k^N = \eta_{\tau_k}^N, \quad k \geq 0.$$

For each site x in Δ, there exists an open path $\gamma = (x = x_0, x_1, \ldots, x_{n(x)})$ such that $x_{n(x)} \in \mathcal{E}$, $R_N(x_i, x_{i+1}) > C_0$, $0 \leq i < n(x)$, for some positive constant C_0, independent of N, whose value may change from line to line. In particular,

$$\mathbf{P}_x^N[Y_k^N = x_k : 0 \leq k \leq n(x)] \geq C_0. \tag{5.4}$$

Let $n = \max\{n(x) : x \in \Delta\}$.

By the strong Markov property and decomposing the space according to the partition $\{T_{\mathcal{E}} \leq \tau_n\}$, $\{T_{\mathcal{E}} > \tau_n\}$, for every $x \in \Delta$, since on the set $\{T_{\mathcal{E}} > \tau_n\}$, $T_{\mathcal{E}} = \tau_n + T_{\mathcal{E}} \circ \tau_n$,

$$\mathbf{E}_x^N[T_{\mathcal{E}}] = \mathbf{E}_x^N[\min\{T_{\mathcal{E}}, \tau_n\}] + \mathbf{E}_x^N[\mathbf{1}\{T_{\mathcal{E}} > \tau_n\}\mathbf{E}_{\eta_{\tau_n}^N}^N[T_{\mathcal{E}}]].$$

As $\eta_{\tau_n}^N$ belongs to Δ when $T_{\mathcal{E}} > \tau_n$, it follows from the previous identity that

$$\max_{x \in \Delta} \mathbf{E}_x^N[T_{\mathcal{E}}] \leq \frac{\displaystyle\max_{x \in \Delta} \mathbf{E}_x^N[\min\{T_{\mathcal{E}}, \tau_n\}]}{1 - \displaystyle\max_{x \in \Delta} \mathbf{P}_x^N[T_{\mathcal{E}} > \tau_n]}.$$

It follows from (5.4) that the denominator is bounded below by a strictly positive constant C_0. To estimate the numerator, observe that $T_{\mathcal{E}} = \sum_{j \geq 1} \tau_j \mathbf{1}\{A_j\}$, where $A_j = \{Y_0^N \in \Delta, \ldots, Y_{j-1}^N \in \Delta, Y_j^N \in \mathcal{E}\}$. Hence, $\min\{T_{\mathcal{E}}, \tau_n\} = \sum_{1 \leq j < n} \tau_j \mathbf{1}\{A_j\} + \tau_n \mathbf{1}\{B_n\}$, where $B_n = \cup_{j \geq n} A_j$. Since on the set A_j, $Y_k^N \in \Delta$, $0 \leq k < j$, on A_j the random time τ_j can be estimated by the sum of j mean C_0 independent exponential random variables. Hence,

$$\max_{x \in \Delta} \mathbf{E}_x^N[\min\{T_{\mathcal{E}}, \tau_n\}] \leq C_0 \sum_{j=1}^{n} j,$$

which concludes the proof of the lemma. $\quad\square$

A similar argument permits us to increase the negligible set $\mathcal{B}_i$ of the valley $(\mathcal{E}_i, \mathcal{E}_i \cup \mathcal{B}_i, x)$.

Lemma 5.4. *Fix* $1 \leq i \leq \nu$ *and* x *in* $\mathcal{E}_i$. *The triple* $(\mathcal{E}_i, \mathcal{E}_i \cup \Delta, x)$ *is a valley of depth* $\theta_{N,i} = \mu_N(\mathcal{E}_i)/\mathrm{cap}_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c)$. *Moreover,* $\mathrm{cap}_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c) \approx \mathrm{cap}_N(\mathcal{E}_i, \breve{\mathcal{E}}_i)$.

Proof. Fix $1 \leq i \leq \nu$ and x in $\mathcal{E}_i$. By Lemmas 5.2 and 10.1, to prove the first claim of the proposition we need to show that for every $\delta > 0$,

$$\lim_{N \to \infty} \max_{y \in \Delta \setminus \mathcal{B}_i} \mathbf{P}_y^N[T_{\breve{\mathcal{E}}_i} > \delta \theta_{N,i}] = 0.$$

Fix y in $\Delta \setminus \mathcal{B}_i$. By definition, $\mu_N(y) \succeq \mu_N(\mathcal{E}_i)$. In particular, there is no open path from y to $\mathcal{E}_i$. Indeed, if $y = y_0, \ldots, y_{n-1} \notin \mathcal{E}_i$, $y_n \in \mathcal{E}_i$ is an open path from y to $\mathcal{E}_i$, the relation $\mu_N(y_{n-1}) \succeq \mu_N(y) \succeq \mu_N(\mathcal{E}_i)$ contradicts the identity $\mu_N(y_{n-1})R_N(y_{n-1}, y_n) = \mu_N(y_n)R_N(y_n, y_{n-1})$ because $\max_{z \in \mathcal{E}_i, z' \notin \mathcal{E}_i} R_N(z, z') \prec 1$.

Recall that we denote by $\{Y_k^N : k \geq 0\}$ the jump chain associated with the Markov process η_t^N. Its jump probabilities $p_N(x, y)$, $x, y \in E$, $x \neq y$, are given by

$$p_N(x, y) = \frac{R_N(x, y)}{\sum\limits_{z \in E} R_N(x, z)}.$$

In view of (2.1), as $N \uparrow \infty$, $p_N(x, y)$ converges to some $p(x, y) \in [0, 1]$ such that $\sum_y p(x, y) = 1$. Let $\{Z_k : k \geq 0\}$ be the discrete time Markov chain associated with the jump probabilities $p(x, y)$. Note that the Markov chain Z_k may not be irreducible.

Clearly, we may couple the two chains in such a way that for every $n \geq 1$,

$$\lim_{N \to \infty} \mathbf{P}_y^N \left[\bigcup_{k=1}^{n} \{Y_k^N \neq Z_k\} \right] = 0. \tag{5.5}$$

On the other hand, before reaching $\mathcal{E}$ the Markov chain Z_k only uses open bonds. Since there is no open path from y to $\mathcal{E}_i$ and since there are open paths from y to $\check{\mathcal{E}}_i$, the chain Z_k eventually reaches $\check{\mathcal{E}}_i$. Hence,

$$\lim_{n \to \infty} \mathbf{P}_y^N \left[\bigcap_{k=1}^{n} \{Z_k \notin \check{\mathcal{E}}_i\} \right] = 0.$$

Recall that $\{\tau_n : n \geq 1\}$ stands for the jump times of the Markov process η_t^N. On the set $[\cap_{k=1}^{n} \{Y_k^N = Z_k\}] \cap [\cup_{k=1}^{n} \{Z_k \in \check{\mathcal{E}}_i\}]$, $T_{\check{\mathcal{E}}_i} \leq \tau_k$ for some $k \leq n$, and τ_k may be bounded by the sum of k mean C_0 i.i.d. exponential random variables, for some finite constant C_0, independent of N. Therefore, since $\theta_{N,i} \succ 1$, for every $n \geq 1$,

$$\lim_{N \to \infty} \mathbf{P}_y^N \left[T_{\check{\mathcal{E}}_i} > \delta\theta_{N,i}, \bigcap_{k=1}^{n} \{Y_k^N = Z_k\}, \bigcup_{k=1}^{n} \{Z_k \in \check{\mathcal{E}}_i\} \right] = 0,$$

which proves the first assertion of the lemma.

In view of Lemma 4.1, to prove the second claim, it is enough to show that $G_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c) \approx G_N(\mathcal{E}_i, \check{\mathcal{E}}_i)$. In fact, we assert that

$$G_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c) = G_N(\mathcal{E}_i, \check{\mathcal{E}}_i) \tag{5.6}$$

for all N sufficiently large.

On the one hand, since $\check{\mathcal{E}}_i \subset [\mathcal{E}_i \cup \mathcal{B}_i]^c$, $G_N(\mathcal{E}_i, \check{\mathcal{E}}_i) \leq G_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c)$. On the other hand, since the set E is finite, there exists a path $\gamma = (x_0, \ldots, x_n)$ in $\Gamma_{\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c}$ such that $G_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c) = G_N(\gamma)$. By definition, $x_0 \in \mathcal{E}_i$, $x_n \notin \mathcal{E}_i \cup \mathcal{B}_i$, and we may assume without loss of generality that $x_1 \notin \mathcal{E}_i$.

Since $x_1 \notin \mathcal{E}_i$ and $\mathcal{E}_i$ is a leaf, $G_N(x_0, x_1) \prec \mu_N(\mathcal{E}_i)$, so

$$G_N(\gamma) = \min_{0 \leq i < n} G_N(x_i, x_{i+1}) \prec \mu_N(\mathcal{E}_i). \tag{5.7}$$

Either x_n belongs to $\Delta \setminus \mathcal{B}_i$ or x_n belongs to $\check{\mathcal{E}}_i$. In the latter case, γ is a path in $\Gamma_{\mathcal{E}_i, \check{\mathcal{E}}_i}$, so $G_N(\gamma) \leq G_N(\mathcal{E}_i, \check{\mathcal{E}}_i)$, proving that $G_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c) \leq G_N(\mathcal{E}_i, \check{\mathcal{E}}_i)$.

If x_n belongs to $\Delta \setminus \mathcal{B}_i$, by the definition of $\mathcal{B}_i$ and the leaves $\mathcal{E}_j$, there exists a self-avoiding path $\tilde{\gamma} = (x_n, x_{n+1}, \ldots, x_l)$ from x_n to $\mathcal{E}$ such that $R_N(x_i, x_{i+1}) \geq C_0$, $n \leq i < l$, for some finite constant C_0, independent of N, whose value may change from line to line. Since

$x_n \in \Delta \setminus \mathcal{B}_i$, $\mu_N(x_n) \geq C_0 \mu_N(\mathcal{E}_i)$ and the same estimate holds for $\mu_N(x_j)$, $n < j \leq l$, because $R_N(x_i, x_{i+1}) \geq C_0$, $n \leq i < l$, and $R_N(y, z) \leq C_0$ for all $y, z \in E$. From these estimates we derive two facts. First, x_l may not belong to $\mathcal{E}_i$ because $\mu_N(x_{l-1}) \geq C_0 \mu_N(\mathcal{E}_i)$, $R_N(x_{l-1}, x_l) \approx 1$ and $R_N(y, z) \prec 1$ for all $y \in \mathcal{E}_i$, $z \in \mathcal{E}_i^c$. Second, $\min_{n \leq i < l} G_N(x_i, x_{i+1}) \geq C_0 \mu_N(\mathcal{E}_i)$ because $R_N(x_i, x_{i+1}) \geq C_0$, $n \leq i < l$.

Therefore, if x_n belongs to $\Delta \setminus \mathcal{B}_i$, juxtaposing the paths γ and $\tilde{\gamma}$, in view of (5.7), we obtain a self-avoiding path from $\mathcal{E}_i$ to $\check{\mathcal{E}}_i$ such that $\min_{1 \leq i < l} G_N(x_i, x_{i+1}) = \min_{1 \leq i < n} G_N(x_i, x_{i+1}) = G_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c)$ for N sufficiently large. Hence, also in the case where x_n belongs to $\Delta \setminus \mathcal{B}_i$, $G_N(\mathcal{E}_i, \check{\mathcal{E}}_i) \geq G_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c)$, which proves (5.6). □

We may in fact compute the asymptotic behavior of the depth $\theta_{N,i}$ of the valley $(\mathcal{E}_i, \mathcal{E}_i \cup \Delta, x)$ with the help of Lemma 4.2. Recall the definition of the Markov chain $\{Z_k : k \geq 0\}$ introduced in the previous proposition. Denote by $\mathbf{P}_x^Z$ the probability on the path space $D(\mathbb{Z}_+, E)$ induced by the Markov chain $\{Z_k : k \geq 0\}$ starting from x.

Lemma 5.5. *Fix a subset I of $\{1, \ldots, \nu\}$ and let $J = \{1, \ldots, \nu\} \setminus I$, $f_N(x) = \mathbf{P}_x^N[T_{\mathcal{E}_I} < T_{\mathcal{E}_J}]$, where $\mathcal{E}_I = \cup_{i \in I} \mathcal{E}_i$. We claim that*

$$\lim_{N \to \infty} \mathbf{P}_x^N[T_{\mathcal{E}_I} < T_{\mathcal{E}_J}] = f_{I,J}(x) := \mathbf{P}_x^Z[T_{\mathcal{E}_I} < T_{\mathcal{E}_J}], \quad x \in E.$$

In particular,

$$\lim_{N \to \infty} \frac{\operatorname{cap}_N(\mathcal{E}_I, \mathcal{E}_J)}{\mathfrak{g}_N(\mathcal{E}_I, \mathcal{E}_J)} = \sum_{(x,y) \in \mathfrak{B}(\mathcal{E}_I, \mathcal{E}_J)} g(x, y)[f_{I,J}(y) - f_{I,J}(x)]^2 \in (0, \infty).$$

Proof. Clearly, for every $x \in E$,

$$\lim_{n \to \infty} \mathbf{P}_x^Z \left[\bigcup_{k=1}^n \{Z_k \in \mathcal{E}\} \right] = 1.$$

It follows from this estimate and from (5.5) that for every $x \in \Delta$,

$$\lim_{N \to \infty} \mathbf{P}_x^N[T_{\mathcal{E}_I} < T_{\mathcal{E}_J}] = \mathbf{P}_x^Z[T_{\mathcal{E}_I} < T_{\mathcal{E}_J}],$$

which proves the first assertion of the lemma. The second one follows from the previous result and Lemma 4.2. □

Note that this result is a particular case of Lemma 4.3. The same argument provides the asymptotic value of the capacity between $\mathcal{E}_i$ and $[\mathcal{E}_i \cup \mathcal{B}_i]^c$.

Lemma 5.6. *Fix $1 \leq i \leq \nu$ and let $f_N : E \to [0, 1]$ be given by $f_N(x) = \mathbf{P}_x^N[T_{\mathcal{E}_i} < T_{[\mathcal{E}_i \cup \mathcal{B}_i]^c}]$. Then, f_N converges pointwise to $f_i(x) = \mathbf{P}_x^Z[T_{\mathcal{E}_i} < T_{[\mathcal{E}_i \cup \mathcal{B}_i]^c}]$. In particular,*

$$\lim_{N \to \infty} \frac{\operatorname{cap}_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c)}{\mathfrak{g}_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c)} = \sum_{(x,y) \in \mathfrak{B}(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c)} g(x, y)[f_i(y) - f_i(x)]^2.$$

We may now state the first main result of this section.

Proposition 5.7. *For $1 \leq i \leq \nu$, $(\mathcal{E}_i, \mathcal{E}_i \cup \Delta, x)$, $x \in \mathcal{E}_i$, is a valley of depth $\theta_{N,i} = \mu_N(\mathcal{E}_i)/\operatorname{cap}_N(\mathcal{E}_i, \check{\mathcal{E}}_i)$. Moreover,*

$$\lim_{N \to \infty} \frac{\operatorname{cap}_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c)}{\operatorname{cap}_N(\mathcal{E}_i, \check{\mathcal{E}}_i)} = 1.$$

Proof. Fix $1 \leq i \leq \nu$ and note that the set $\mathcal{E}_J$ appearing in Lemma 5.5 is equal to $\check{\mathcal{E}}_i$ if $I = \{i\}$. By (5.6), $\mathfrak{g}_N(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c) = \mathfrak{g}_N(\mathcal{E}_i, \check{\mathcal{E}}_i)$, so $\mathfrak{B}(\mathcal{E}_i, [\mathcal{E}_i \cup \mathcal{B}_i]^c) = \mathfrak{B}(\mathcal{E}_i, \check{\mathcal{E}}_i)$. Let $g_i = f_{I,J}$ when $I = \{i\}$. Since there is an open path from any state in $\Delta \setminus \mathcal{B}_i$ to $\check{\mathcal{E}}_i$ and no open path from a state in $\Delta \setminus \mathcal{B}_i$ to $\mathcal{E}_i$, $f_i(x) = \mathbf{P}_x^Z[T_{\mathcal{E}_i} < T_{[\mathcal{E}_i \cup \mathcal{B}_i]^c}] = \mathbf{P}_x^Z[T_{\mathcal{E}_i} < T_{\check{\mathcal{E}}_i}] = g_i(x)$. This proves the corollary in view of Lemmas 5.5 and 5.6. □

By Proposition 5.7 and Lemma 5.5,

$$u_i := \lim_{N \to \infty} \frac{\mathfrak{g}_N(\mathcal{E}_i, \check{\mathcal{E}}_i)}{\mu_N(\mathcal{E}_i)} \theta_{N,i} \in (0, \infty). \tag{5.8}$$

In particular, the depths of the valleys are comparable.

Proposition 5.8. *The sequences* $(\theta_{N,i} : N \geq 1)$, $1 \leq i \leq \nu$, *are comparable and* $\theta_{N,i} \succ 1$, $1 \leq i \leq \nu$.

Proof. To prove this lemma, we have to show that, as $N \uparrow \infty$, the sequences $\theta_{N,i}/\theta_{N,j}, i \neq j$, vanish, diverge, or converge. Fix $i \neq j$. By (5.8),

$$\lim_{N \to \infty} \frac{\theta_{N,i}}{\theta_{N,j}} = \frac{u_i}{u_j} \lim_{N \to \infty} \frac{\mu_N(\mathcal{E}_i)}{\mu_N(\mathcal{E}_j)} \frac{\mathfrak{g}_N(\mathcal{E}_j, \check{\mathcal{E}}_j)}{\mathfrak{g}_N(\mathcal{E}_i, \check{\mathcal{E}}_i)}.$$

By (5.2) and (2.4), $\mu_N(\mathcal{E}_k) = \mu_N(x_k)a_N$ for some $x_k \in \mathcal{E}_k$ and some sequence a_N which converges to some $a \in (0, \infty)$ as $N \uparrow \infty$. On the other hand, by definition, $\mathfrak{g}_N(\mathcal{E}_k, \check{\mathcal{E}}_k) = G_N(y_k, z_k)b_N = \mu_N(y_k)R_N(y_k, z_k)b_N$ for some bond (y_k, z_k), where b_N converges to some limit $b \in (0, \infty)$ as $N \uparrow \infty$. Hence,

$$\frac{\theta_{N,i}}{\theta_{N,j}} = c_N \frac{\mu_N(x_i)\,\mu_N(y_j)\,R_N(y_j, z_j)}{\mu_N(x_j)\,\mu_N(y_i)\,R_N(y_i, z_i)}$$

for some sequence c_N which converges to some limit $c \in (0, \infty)$ as $N \uparrow \infty$. In view of identity (2.3) and assumption (2.2), the sequences $\theta_{N,i}$ are comparable. This proves the first assertion of the lemma.

Fix $1 \leq i \leq \nu$ and recall the definition of $\theta_{N,i}$ given in Proposition 5.7. By Lemma 4.1, it is enough to show that $\mu_N(\mathcal{E}_i)/G_N(\mathcal{E}_i, \check{\mathcal{E}}_i) \succ 1$. Fix a self-avoiding path $\gamma = (x_0, \ldots, x_n)$ from $\mathcal{E}_i$ to $\check{\mathcal{E}}_i$ such that $G_N(\gamma) = G_N(\mathcal{E}_i, \check{\mathcal{E}}_i)$. There exists a bond (x_j, x_{j+1}) such that $x_j \in \mathcal{E}_i, x_{j+1} \notin \mathcal{E}_i$. By the definition of $G_N(\gamma)$, by (5.2) and since $\mathcal{E}_i$ is a leaf, $G_N(\gamma) \leq G_N(x_j, x_{j+1}) \prec \mu_N(x_j) \approx \mu_N(\mathcal{E}_i)$, proving the second assertion of the lemma. □

6. Metastability among the shallowest valleys

We describe in this section the asymptotic behavior of the Markov process $\{\eta_t^N : t \geq 0\}$ on the smallest time scale needed for the process to jump from one leaf to another.

Let $\theta_N(1) = \min\{\theta_{N,i} : 1 \leq i \leq \nu\}$ and denote by S_1 the indices of the shallowest leaves, i.e., the ones whose valleys have depth of magnitude $\theta_N(1)$:

$$S_1 = \{i : \theta_{N,i} \approx \theta_N(1)\}.$$

Since, by Proposition 5.8, the depths of the valleys are comparable and since $\theta_N(1)$ is the depth of the shallowest valley, $\theta_N(1)/\theta_{N,i}$ converges as $N \uparrow \infty$:

$$\lambda(i) := \lim_{N \to \infty} \frac{\theta_N(1)}{\theta_{N,i}} \in (0, \infty). \tag{6.1}$$

Lemma 6.1. *For any $1 \le i \ne j \le \nu$, $\theta_N(1)\, r_N^{\mathcal{E}}(\mathcal{E}_i, \mathcal{E}_j)$ converges, as $N \uparrow \infty$, to some number $r(i, j) \in [0, \infty)$.*

Proof. Fix $1 \le i \ne j \le \nu$. By [2, Lemma 6.7] and by Proposition 5.7, we may rewrite $\theta_N(1)\, r_N^{\mathcal{E}}(\mathcal{E}_i, \mathcal{E}_j)$ as

$$\theta_N(1)\, r_N^{\mathcal{E}}(\mathcal{E}_i, \check{\mathcal{E}}_i) \frac{r_N^{\mathcal{E}}(\mathcal{E}_i, \mathcal{E}_j)}{r_N^{\mathcal{E}}(\mathcal{E}_i, \check{\mathcal{E}}_i)} = \theta_N(1) \frac{\mathrm{cap}_N(\mathcal{E}_i, \check{\mathcal{E}}_i)}{\mu_N(\mathcal{E}_i)} \frac{r_N^{\mathcal{E}}(\mathcal{E}_i, \mathcal{E}_j)}{r_N^{\mathcal{E}}(\mathcal{E}_i, \check{\mathcal{E}}_i)} = \frac{\theta_N(1)}{\theta_{N,i}} \frac{r_N^{\mathcal{E}}(\mathcal{E}_i, \mathcal{E}_j)}{r_N^{\mathcal{E}}(\mathcal{E}_i, \check{\mathcal{E}}_i)}.$$

By (6.1), $\theta_N(1)/\theta_{N,i}$ converges to $\lambda(i)$. On the other hand, by Corollary 4.4, $r_N^{\mathcal{E}}(\mathcal{E}_i, \mathcal{E}_j)/r_N^{\mathcal{E}}(\mathcal{E}_i, \check{\mathcal{E}}_i)$ converges, as $N \uparrow \infty$, to some number $q(i, j) \in [0, 1]$. This proves the lemma with $r(i, j) = \lambda(i)\, q(i, j)$. $\square$

Let $\Psi : \mathcal{E} \to \{1, \ldots, \nu\}$ be given by $\Psi(x) = \sum_{1 \le i \le \nu} i\, \mathbf{1}\{x \in \mathcal{E}_i\}$.

Lemma 6.2. *Fix $1 \le i \le \nu$ and $x \in \mathcal{E}_i$. Under $\mathbf{P}_x^N$, the speeded-up process $X_t^N = \Psi(\eta_{t\theta_N(1)}^{\mathcal{E}})$ converges to a Markov process on $\{1, \ldots, \nu\}$ with rates $r(\cdot, \cdot)$ starting from i.*

Proof. We need to check that the assumptions of [2, Theorem 2.7] are fulfilled. On the one hand, condition (H1) follows from Lemma 5.1 and Proposition 5.8 which asserts that $\theta_{N,i} \uparrow \infty$ as $N \uparrow \infty$. On the other hand, condition (H0) has been proven in Lemma 6.1. $\square$

Note that $\lambda(j) = 0$ if $j \notin S_1$. The points in S_1^c are therefore absorbing for the asymptotic dynamics.

Recall Definition 3.7 of [2]. The main result of this section, stated below in Proposition 6.3, asserts that the Markov process $\{\eta_t^N : t \ge 0\}$ exhibits a metastable behavior on the time scale $\theta_N(1)$ with asymptotic dynamics characterized by the jump rates $r(i, j)$ introduced in Lemma 6.1. Denote by $\{\mathbb{P}_i : 1 \le i \le \nu\}$ the laws on the path space $D(\mathbb{R}_+, \{1, \ldots, \nu\})$ of a Markov process on $\{1, \ldots, \nu\}$ whose sites in S_1^c are absorbing and which jumps from $i \in S_1$ to $j \ne i$ at rate $r(i, j)$.

Proposition 6.3. *Fix a site x_i on each leaf $\mathcal{E}_i$. The sequence of Markov processes $\{\eta_t^N : t \ge 0\}$ exhibits a metastable behavior on the time scale $\theta_N(1)$ with metastates $\{\mathcal{E}_i : 1 \le i \le \nu\}$, metapoints $\{x_i : 1 \le i \le \nu\}$ and asymptotic Markov dynamics $\{\mathbb{P}_i : 1 \le i \le \nu\}$.*

Proof. Condition (M2) has been proven in Lemma 6.2.

To prove (M3$'$), observe that for every $x \in E$,

$$\mathbf{E}_x^N\left[\int_0^t \mathbf{1}\{\eta_{s\theta_N(1)}^N \in \Delta\}\, \mathrm{d}s\right] \le \max_{y \in \Delta} \mathbf{E}_y^N\left[\int_0^t \mathbf{1}\{\eta_{s\theta_N(1)}^N \in \Delta\}\, \mathrm{d}s\right].$$

Fix $y \in \Delta$ and denote by $U_k, V_k, k \ge 1$, the successive lengths of the sojourns in Δ and Δ^c:

$$U_1 = \inf\{t > 0 : \eta_t^N \notin \Delta\}, \qquad V_1 = \inf\{t > 0 : \eta_{t+U_1}^N \in \Delta\},$$

$$U_{k+1} = \inf\{t > 0 : \eta_{t+V_k}^N \notin \Delta\}, \qquad V_{k+1} = \inf\{t > 0 : \eta_{t+U_{k+1}}^N \in \Delta\}.$$

Denote by $\{N_t : t \ge 0\}$ the counting process associated with the sequence $\{V_k : k \ge 1\}$: $\{N_t = k\} = \{V_1 + \cdots + V_k \le t < V_1 + \cdots + V_{k+1}\}$, $k \ge 0$, and observe that

$$\int_0^{t\theta_N(1)} \mathbf{1}\{\eta_s^N \in \Delta\}\, \mathrm{d}s \le U_1 + \sum_{k=1}^{N_{t\theta_N(1)}} U_{k+1}.$$

Let $\lambda_N = \max_{1 \le j \le v} \max_{y \in \mathcal{E}_j} \sum_{z \notin \mathcal{E}_j} R_N(y, z) \prec 1$. We may estimate from below the random variables $\{V_k : k \ge 1\}$ by independent exponential times of rate $\lambda_N : V_k \ge \hat{V}_k$, where $\{\hat{V}_k : k \ge 1\}$ is a sequence of i.i.d. mean λ_N^{-1} exponential random variables, independent also from the sequence $\{U_k : k \ge 1\}$.

Let $\{\hat{N}_t : t \ge 0\}$ be the Poisson process associated with the sequence $\{\hat{V}_k : k \ge 1\}$. Since the sequence $\{\hat{V}_k : k \ge 1\}$ is independent of the sequence $\{U_k : k \ge 1\}$, in view of the previous estimate,

$$\mathbf{E}_y^N \left[\int_0^{t\theta_N(1)} \mathbf{1}\{\eta_s^N \in \Delta\} \, \mathrm{d}s \right] \le \mathbf{E}_y^N \left[U_1 + \sum_{k=1}^{\hat{N}_{t\theta_N(1)}} U_{k+1} \right]$$

$$\le \mathbf{E}_y^N[U_1] + \sum_{\ell \ge 0} \mathbf{E}_y^N \left[\sum_{k=1}^{\ell} U_{k+1} \right] P_y^N \left[\hat{N}_{t\theta_N(1)} = \ell \right].$$

By Lemma 5.3 this expression is bounded by $C_0\{1 + t\,\theta_N(1)\,\lambda_N\}$, which proves condition (M3′).

The proof of condition (M1′) is similar to that of Lemma 5.3. However, we may not estimate the expectation of T_{x_i} which might be very large if the process leaves the metastable set $\mathcal{E}_i$ before reaching the state x_i. We may of course assume that $\mathcal{E}_i$ is not a singleton, so $\sum_{z \in E} R_N(y, z)$ is of magnitude 1 for all $y \in \mathcal{E}_i$.

Recall that we denote by $\{\tau_k : k \ge 0\}$ the successive jump times of η_t^N and by $\{Y_k^N : k \ge 0\}$ the jump chain. Fix $1 \le i \le v$, $x_i \in \mathcal{E}_i$ and let now $\lambda_N = \max_{y \in \mathcal{E}_i, z \notin \mathcal{E}_i} R_N(y, z) \prec 1$. For each $y \in \mathcal{E}_i$, there exists an open path $\gamma = (y_0 = y, \dots, y_{n(y)} = x)$ from y to x contained in $\mathcal{E}_i$. Let $n = \max\{n(y) : y \in \mathcal{E}_i, y \ne x\}$. There exists a constant a, independent of N, such that

$$\max_{y \in \mathcal{E}_i} \mathbf{P}_y^N \left[Y_k^N \ne x_i, \, 0 \le k \le n \right] \le a < 1.$$

On the one hand, for every $\ell \ge 1$, $y \in \mathcal{E}_i$,

$$\mathbf{P}_y^N \left[\tau_{\ell n} \le \min\{T_{x_i}, T_{\mathcal{E}_i^c}\} \right] \le \mathbf{P}_y^N \left[Y_k^N \ne x_i, Y_k^N \in \mathcal{E}_i, 0 \le k \le \ell n \right].$$

By the Markov property and by the previous estimate, this expression is bounded by a^ℓ. On the other hand, since the process jumps from $\mathcal{E}_i$ to $\mathcal{E}_i^c$ at rate $\lambda_N \prec 1$,

$$\mathbf{P}_y^N \left[\tau_{n\ell} \ge T_{\mathcal{E}_i^c} \right] \le \mathbf{P}_y^N \left[\bigcup_{k=1}^{n\ell} Y_k^N \notin \mathcal{E}_i \right] \le C_0 \ell n \lambda_N$$

for some finite constant C_0 independent of N.

In view of the previous bounds, to estimate $\mathbf{P}_y^N[T_{x_i} > \delta\theta_N]$ it remains to consider the term

$$\mathbf{P}_y^N \left[T_{x_i} > \delta\theta_N, T_{x_i} < \tau_{n\ell} < T_{\mathcal{E}_i^c} \right] \le \mathbf{P}_y^N \left[\delta\theta_N < \tau_{n\ell} < T_{\mathcal{E}_i^c} \right].$$

Since $\sum_{z \in E} R_N(y, z)$ is of magnitude 1 for all $y \in \mathcal{E}_i$, before leaving the set $\mathcal{E}_i$, we may estimate the times between jumps by i.i.d. exponential random times with finite mean independent of N. By the Tchebycheff inequality, the previous expression is thus bounded by $C_0 n\ell / \delta\theta_N$ for some finite constant C_0 independent of N.

We have thus proved that for every $\delta > 0$, $y \in \mathcal{E}_i$,

$$\mathbf{P}_y^N \left[T_{x_i} > \delta\theta_N \right] \le a^\ell + C_0 \, \ell \, n \, \lambda_N + \frac{C_0 n\ell}{\delta\theta_N}.$$

The second assertion of the lemma follows by taking $N \uparrow \infty$ and then $\ell \uparrow \infty$. □

We conclude this section with two remarks. Denote by $P_N(x, i, j)$, $1 \leq i \neq j \leq \nu$, $x \in \mathcal{E}_i$, the hitting probabilities

$$P_N(x, i, j) := \mathbf{P}_x^N [T_{\mathcal{E}_j} = T_{\check{\mathcal{E}}_i}].$$

By Lemma 4.3, $P_N(x, i, j)$ converges to some $P(x, i, j) \in [0, 1]$. Since, by Proposition 5.7, $(\mathcal{E}_i, \mathcal{E}_i \cup \Delta, y)$, $y \in \mathcal{E}_i$, is a valley, it is not difficult to show that the limit $P(x, i, j)$ does not depend on the starting point x. Therefore, by Lemma 10.2, for any $1 \leq i \neq j \leq \nu$,

$$r(i, j) = \lambda(i) \, p(i, j), \quad \text{where } p(i, j) := \lim_{N \to \infty} \mathbf{P}_x^N [T_{\mathcal{E}_j} = T_{\check{\mathcal{E}}_i}] \tag{6.2}$$

and where $\lambda(i)$ is defined in (6.1).

Consider a leaf $\mathcal{E}_i$, $i \in S_1$, and a leaf $\mathcal{E}_j$ such that $\mu_N(\mathcal{E}_j) \prec \mu_N(\mathcal{E}_i)$. By reversibility,

$$\mu_N(\mathcal{E}_i) \, r_N^{\mathcal{E}}(\mathcal{E}_i, \mathcal{E}_j) = \mu_N(\mathcal{E}_j) \, r_N^{\mathcal{E}}(\mathcal{E}_j, \mathcal{E}_i).$$

By [2, Lemma 6.7], $r_N^{\mathcal{E}}(\mathcal{E}_j, \mathcal{E}_i) \leq r_N^{\mathcal{E}}(\mathcal{E}_j, \check{\mathcal{E}}_j) = \text{cap}_N(\mathcal{E}_j, \check{\mathcal{E}}_j)/\mu_N(\mathcal{E}_j) = 1/\theta_{N,j}$, so $\theta_N(1) \, r_N^{\mathcal{E}}(\mathcal{E}_j, \mathcal{E}_i)$ is bounded. Therefore, $\theta_N(1) \, r_N^{\mathcal{E}}(\mathcal{E}_i, \mathcal{E}_j)$ vanishes as $N \uparrow \infty$. We have just proved that

$$r(i, j) = \lim_{N \to \infty} \theta_N(1) \, r_N^{\mathcal{E}}(\mathcal{E}_i, \mathcal{E}_j) = 0 \quad \text{for all } j; \ \mu_N(\mathcal{E}_j) \prec \mu_N(\mathcal{E}_i). \tag{6.3}$$

Hence, in the asymptotic dynamics, the process may only jump from a leaf $\mathcal{E}_i$ to a leaf $\mathcal{E}_j$ if the measure of $\mathcal{E}_j$ is of the same magnitude as or of a larger magnitude than that of $\mathcal{E}_i$.

7. Multi-scale analysis

In the previous section, we proved that the Markov process $\{\eta_t^N : t \geq 0\}$ exhibits a metastable behavior on the time scale $\theta_N(1)$ with metastates $\{\mathcal{E}_i : 1 \leq i \leq \nu\}$, metapoints $\{x_i : 1 \leq i \leq \nu\}$ and asymptotic Markov dynamics $\{\mathbb{P}_i : 1 \leq i \leq \nu\}$.

We describe in this section, by a recursive argument, the metastable behavior of the Markov process $\{\eta_t^N : t \geq 0\}$ on longer time scales. In the statement of the hypothesis T below, by convention, $\theta_N(0) \equiv 1$, $\nu(0) = |E|$ and the sets $\mathcal{E}_1^{(0)}, \ldots, \mathcal{E}_{\nu(0)}^{(0)}$ are all singletons of E.

Assumption T *at level* $\mathfrak{p}$: For each $1 \leq k \leq \mathfrak{p}$ there exists a sequence $(\theta_N(k) : N \geq 1)$, $1 \prec \theta_N(k) \prec \theta_N(k+1)$, $1 \leq k < \mathfrak{p}$, and a partition $\{\mathcal{E}_1^{(k)}, \ldots, \mathcal{E}_{\nu(k)}^{(k)}, \Delta_k\}$ of the state space E, such that:

(T1) $1 \leq \nu(k) < \nu(k-1)$.
(T2) For $1 \leq i \leq \nu(k)$, $\mathcal{E}_i^{(k)} = \cup_{j \in I_{k,i}} \mathcal{E}_j^{(k-1)}$, where $I_{k,1}, \ldots, I_{k,\nu(k)}$ are disjoint subsets of $\{1, \ldots, \nu(k-1)\}$.
(T3) For all $1 \leq i \leq \nu(k)$, $\mu_N(x) \approx \mu_N(\mathcal{E}_i^{(k)})$ for all $x \in \mathcal{E}_i^{(k)}$.
(T4) There exists a positive constant C_1, independent of N, such that for all $1 \leq i \leq \nu(k)$ and all $x, y \in \mathcal{E}_i^{(k)}$, $x \neq y$, $\text{cap}_N(x, y) \geq C_1 \mu_N(\mathcal{E}_i^{(k)})/\theta_N(k-1)$.
(T5) For all $1 \leq i \leq \nu(k)$, $\mu_N(\mathcal{E}_i^{(k)})/\text{cap}_N(\mathcal{E}_i^{(k)}, \check{\mathcal{E}}_i^{(k)}) \succeq \theta_N(k)$, where $\check{\mathcal{E}}_i^{(k)} = \cup_{j \neq i} \mathcal{E}_j^{(k)}$.
(T6) Let

$$\mathcal{E}^{(k)} = \bigcup_{i=1}^{\nu(k)} \mathcal{E}_i^{(k)}, \qquad S_k = \left\{ i : \frac{\mu_N(\mathcal{E}_i^{(k)})}{\text{cap}_N(\mathcal{E}_i^{(k)}, \check{\mathcal{E}}_i^{(k)})} \approx \theta_N(k) \right\}.$$

Then,

$$\lim_{N\to\infty} \theta_N(k)\, r_N^{\mathcal{E}^{(k)}}(\mathcal{E}_i^{(k)}, \mathcal{E}_j^{(k)}) = \mathfrak{r}_k(i,j), \quad 1 \le i \ne j \le \nu(k),$$

$$\sum_{j\ne i} \mathfrak{r}_k(i,j) > 0 \text{ for each } i \in S_k \text{ and } \sum_{j\ne i} \mathfrak{r}_k(i,j) = 0 \text{ for each } i \notin S_k, \tag{7.1}$$

$$\text{and} \quad \mathfrak{r}_k(i,j) = 0 \quad \text{if } \mu_N(\mathcal{E}_j^{(k)}) \prec \mu_N(\mathcal{E}_i^{(k)}).$$

Moreover, recall the definition of the speeded-up blind process $X_t^{N,k} = \Psi_k(\eta_{t\theta_N(k)}^{N,k})$ introduced in the statement of Theorem 2.1. For every $1 \le i \le \nu(k)$, and $x \in \mathcal{E}_i^{(k)}$, under the measure $\mathbf{P}_x^N$,

the speeded-up blind process $X_t^{N,k}$ converges $\hspace{3cm}$ (7.2)

to a Markov process on $\{1, \dots, \nu(k)\}$ characterized by rates $\mathfrak{r}_k(l,m), 1 \le l \ne m \le \nu(k)$, starting from i.

(T7) Property (M1′) of metastability holds. For every $1 \le i \le \nu(k)$, every $x \in \mathcal{E}_i^{(k)}$ and $\delta > 0$,

$$\lim_{N\to\infty} \max_{y\in\mathcal{E}_i^{(k)}} \mathbf{P}_y^N\big[T_x > \delta\theta_N(k)\big] = 0.$$

(T8) Property (M3′) of metastability holds. For every $t > 0$,

$$\lim_{N\to\infty} \max_{x\in E} \mathbf{E}_x^N\left[\int_0^t \mathbf{1}\{\eta_{s\theta_N(k)}^N \in \Delta_k\}\,ds\right] = 0.$$

Note that all these properties have been proved in the previous section for $\mathfrak{p} = 1$ with: $\nu(1) = \nu$; $\mathcal{E}_1^{(1)}, \dots, \mathcal{E}_{\nu(1)}^{(1)}, \Delta_1$ given by the sets $\mathcal{E}_1, \dots, \mathcal{E}_\nu, \Delta$ defined just before (5.1); $\theta_N(1) = \theta_N = p_N(1)$ defined at the beginning of Section 6; and $\mathfrak{r}_1 = r$ defined in Lemma 6.1.

The main result of this section states that if assumption T holds at level $\mathfrak{p}$ and $\nu(\mathfrak{p}) \ge 2$, then it holds at level $\mathfrak{p} + 1$.

To begin the recursive argument, suppose that $\nu(\mathfrak{p}) > 1$. We first describe the metastates at level $\mathfrak{p} + 1$. We say that there exists an open path from $\mathcal{E}_a^{(\mathfrak{p})}$ to $\mathcal{E}_b^{(\mathfrak{p})}$ if there exists a sequence $a = a_0, a_1, \dots, a_n = b$ such that $\mathfrak{r}_\mathfrak{p}(a_k, a_{k+1}) > 0$, where $\mathfrak{r}_\mathfrak{p}$ is the asymptotic jump rate introduced in (7.2). We say that two sets $\mathcal{E}_a^{(\mathfrak{p})}, \mathcal{E}_b^{(\mathfrak{p})}$ are equivalent, $\mathcal{E}_a^{(\mathfrak{p})} \sim \mathcal{E}_b^{(\mathfrak{p})}$, if there exist an open path from $\mathcal{E}_a^{(\mathfrak{p})}$ to $\mathcal{E}_b^{(\mathfrak{p})}$ and an open path from $\mathcal{E}_b^{(\mathfrak{p})}$ to $\mathcal{E}_a^{(\mathfrak{p})}$.

Two equivalent sets $\mathcal{E}_a^{(\mathfrak{p})}, \mathcal{E}_b^{(\mathfrak{p})}$ have measure of the same magnitude. Indeed, if $\mathcal{E}_a^{(\mathfrak{p})}, \mathcal{E}_b^{(\mathfrak{p})}$ are equivalent, there exists an open path $(a = a_0, \dots, a_n = b, \dots, a_{n+m} = a)$ from $\mathcal{E}_a^{(\mathfrak{p})}$ to $\mathcal{E}_a^{(\mathfrak{p})}$ passing through $\mathcal{E}_b^{(\mathfrak{p})}$. By (7.1), $\mu_N(\mathcal{E}_{a_i}^{(\mathfrak{p})}) \preceq \mu_N(\mathcal{E}_{a_{i+1}}^{(\mathfrak{p})})$, $0 \le i < n + m$. Since $\mathcal{E}_{a_0}^{(\mathfrak{p})} = \mathcal{E}_{a_{n+m}}^{(\mathfrak{p})} = \mathcal{E}_a^{(\mathfrak{p})}$, we obtain that

$$\mu_N(\mathcal{E}_a^{(\mathfrak{p})}) \approx \mu_N(\mathcal{E}_b^{(\mathfrak{p})}), \tag{7.3}$$

as claimed.

We call a metastate in the time scale $\theta_N(k)$ a k-metastate. If we declare a $\mathfrak{p}$-metastate equivalent to itself, the relation $\sim$ introduced in the penultimate paragraph becomes an equivalence relation among the $\mathfrak{p}$-metastates $\mathcal{E}_1^{(\mathfrak{p})}, \dots, \mathcal{E}_{\nu(\mathfrak{p})}^{(\mathfrak{p})}$. Denote by $\mathcal{C}_1^{(\mathfrak{p}+1)}, \mathcal{C}_2^{(\mathfrak{p}+1)}, \dots, \mathcal{C}_{\alpha(\mathfrak{p}+1)}^{(\mathfrak{p}+1)}$ the equivalence classes. Some equivalence classes are connected to other equivalence classes.

On drawing an arrow from a set $\mathcal{C}_i^{(\mathfrak{p}+1)}$ to a set $\mathcal{C}_j^{(\mathfrak{p}+1)}$ if there exist $\mathcal{E}_a^{(\mathfrak{p})} \subset \mathcal{C}_i^{(\mathfrak{p}+1)}$, $\mathcal{E}_b^{(\mathfrak{p})} \subset \mathcal{C}_j^{(\mathfrak{p}+1)}$ such that $\mathfrak{r}_\mathfrak{p}(a, b) > 0$, the set $\{\mathcal{C}_1^{(\mathfrak{p}+1)}, \ldots, \mathcal{C}_{\alpha(\mathfrak{p}+1)}^{(\mathfrak{p}+1)}\}$ becomes an oriented graph with no directed loops. We denote by $\mathcal{E}_1^{(\mathfrak{p}+1)}, \mathcal{E}_2^{(\mathfrak{p}+1)}, \ldots, \mathcal{E}_{\nu(\mathfrak{p}+1)}^{(\mathfrak{p}+1)}$ the leaves of this graph, i.e., the set of equivalence classes with no successors in the terminology of graph theory, and by $\Delta_{\mathfrak{p}+1}^o$ the union of the remaining sets such that $\{\mathcal{E}_1^{(\mathfrak{p}+1)}, \ldots, \mathcal{E}_{\nu(\mathfrak{p}+1)}^{(\mathfrak{p}+1)}, \Delta_{\mathfrak{p}+1}\}$, $\Delta_{\mathfrak{p}+1} = \Delta_{\mathfrak{p}+1}^o \cup \Delta_\mathfrak{p}$, forms a partition of E:

$$E = \mathcal{E}^{(\mathfrak{p}+1)} \cup \Delta_{\mathfrak{p}+1}, \qquad \mathcal{E}^{(\mathfrak{p}+1)} = \mathcal{E}_1^{(\mathfrak{p}+1)} \cup \cdots \cup \mathcal{E}_{\nu(\mathfrak{p}+1)}^{(\mathfrak{p}+1)}.$$

For $1 \le i \le \nu(\mathfrak{p}+1)$, let $\breve{\mathcal{E}}_i^{(\mathfrak{p}+1)}$ be the union of all leaves except $\mathcal{E}_i^{(\mathfrak{p}+1)}$:

$$\breve{\mathcal{E}}_i^{(\mathfrak{p}+1)} = \bigcup_{j \ne i} \mathcal{E}_j^{(\mathfrak{p}+1)}.$$

We may now state the main result of this section.

Theorem 7.1. *Let $\{\eta_t^N : t \ge 0\}$ be a sequence of irreducible, reversible Markov processes on a finite state space E satisfying assumptions* (2.1) *and* (2.2). *Suppose that assumption* T *at level $\mathfrak{p}$ holds and that $\nu(\mathfrak{p}) \ge 2$. Define $\nu(\mathfrak{p}+1)$, $\mathcal{E}_i^{(\mathfrak{p}+1)}$, $\breve{\mathcal{E}}_i^{(\mathfrak{p}+1)}$, $1 \le i \le \nu(\mathfrak{p}+1)$, $\mathcal{E}^{(\mathfrak{p}+1)}$, $\Delta_{\mathfrak{p}+1}^o$, $\Delta_{\mathfrak{p}+1}$ as above. Then:*

(1) *For $1 \le i \le \nu(\mathfrak{p}+1)$ and x in $\mathcal{E}_i^{(\mathfrak{p}+1)}$, the triple $(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{E}_i^{(\mathfrak{p}+1)} \cup \Delta_{\mathfrak{p}+1}^o, x)$ is a valley for the trace process $\{\eta_t^{N,\mathfrak{p}} : t \ge 0\}$ of depth $\theta_{N,i} = \mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})/\mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})$.*
(2) *The sequences $(\theta_{N,i} : N \ge 1)$, $1 \le i \le \nu(\mathfrak{p}+1)$ are comparable.*
(3) *Let $\theta_N(\mathfrak{p}+1) = \min\{\theta_{N,i} : 1 \le i \le \nu(\mathfrak{p}+1)\}$. Then, $\theta_N(\mathfrak{p}+1) \succ \theta_N(\mathfrak{p})$.*
(4) *Assumption* T *at level $\mathfrak{p}+1$ holds.*

In the next remark, we summarize what information is needed in each model to prove all of its metastable behavior. It says, in essence, that to prove the metastable behavior of a particular dynamics, we need only to obtain information on the measure, on the capacity and on the times of hitting of subsets of the process.

Remark 7.2. In the applications, once the metastable behavior in the time scale $\theta_N(1)$ among the shallowest valleys has been determined, we shall use Theorem 7.1 to describe the metastable behavior of the process in the longer time scales. We first characterize the k-metastates following the recipe presented above the statement of Theorem 7.1. According to this theorem, the k-metastates form valleys of different depths. To determine the time scale at which metastability at level k can be observed we need to compute the depth of each valley. This computation requires estimates on the capacities among metastates and estimates on the measure of each metastate. Once this has been done, we may define the time scale $\theta_N(k)$. At this point, to complete the description of the metastable behavior of the process at level k, it remains to obtain the rates $\mathfrak{r}_k(i, j)$. Theorem 7.1 asserts that the asymptotic rates $\mathfrak{r}_k(i, j)$ exist. By (7.9) the rates may be expressed in terms of the asymptotic depths of the valleys and the hitting probabilities of the metastates. Hence, to conclude we need to compute, in each model, the limit of the hitting probabilities defined in (7.8), which exist by virtue of Lemma 4.3.

For some evolutions, such as the Kawasaki dynamics, it may be difficult to obtain an exact expression for the limit of the hitting probabilities. Nevertheless, if we may at least determine whether the rates $\mathfrak{r}_k(i, j)$ are positive or equal to 0, we may apply Theorem 7.1 and determine

the time scales at which a metastable behavior is observed and the metastates at each time scale, without an exact knowledge of the asymptotic dynamics among the metastates.

The proof of Theorem 7.1 is divided into several lemmas. We first show that conditions (T1) and (T2) are satisfied for $k = \mathfrak{p} + 1$.

Lemma 7.3. *We have that $\nu(\mathfrak{p} + 1) < \nu(\mathfrak{p})$ and that $\mathcal{E}_i^{(\mathfrak{p}+1)} = \cup_{a \in I_{\mathfrak{p}+1,i}} \mathcal{E}_a^{(\mathfrak{p})}$, $1 \leq i \leq \nu(\mathfrak{p} + 1)$, where $I_{\mathfrak{p}+1,1}, \ldots, I_{\mathfrak{p}+1,\nu(\mathfrak{p}+1)}$ are disjoint subsets of $\{1, \ldots, \nu(\mathfrak{p})\}$.*

Proof. A $\mathfrak{p}$-metastate $\mathcal{E}_a^{(\mathfrak{p})}$, $a \in S_{\mathfrak{p}}$, is either contained in $\Delta_{\mathfrak{p}+1}^o$ or part of a larger leaf $\mathcal{E}_i^{(\mathfrak{p}+1)}$, in the sense that $\mathcal{E}_a^{(\mathfrak{p})} \subsetneq \mathcal{E}_i^{(\mathfrak{p}+1)}$, because by (7.1) each $\mathfrak{p}$-metastate whose index belongs to $S_{\mathfrak{p}}$ has at least one successor. In particular, the number of leaves at level $\mathfrak{p} + 1$ is strictly smaller than the number of $\mathfrak{p}$-metastates, so $\nu(\mathfrak{p} + 1) < \nu(\mathfrak{p})$, proving condition (T1). Condition (T2) follows from the construction. $\square$

Next lemma shows that conditions (T3), (T4) are in force for $k = \mathfrak{p} + 1$.

Lemma 7.4. *For all $1 \leq i \leq \nu(\mathfrak{p} + 1)$, $x \in \mathcal{E}_i^{(\mathfrak{p}+1)}$, $\mu_N(x) \approx \mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})$. Moreover, there exists a positive constant C_1 such that for all $1 \leq i \leq \nu(\mathfrak{p} + 1)$ and all $x, y \in \mathcal{E}_i^{(\mathfrak{p}+1)}$, $x \neq y$, $\mathrm{cap}_N(x, y) \geq C_1 \mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})/\theta_N(\mathfrak{p})$.*

Proof. Fix $1 \leq i \leq \nu(\mathfrak{p} + 1)$, $x \in \mathcal{E}_i^{(\mathfrak{p}+1)}$. By definition, the leaf $\mathcal{E}_i^{(\mathfrak{p}+1)}$ is the union of $\mathfrak{p}$-metastates: $\mathcal{E}_i^{(\mathfrak{p}+1)} = \cup_{a \in I} \mathcal{E}_a^{(\mathfrak{p})}$, where I is a subset of $\{1, \ldots, \nu(\mathfrak{p})\}$.

By (7.3), all $\mathfrak{p}$-metastates $\mathcal{E}_a^{(\mathfrak{p})}$, $a \in I$, have measures of the same magnitude, so $\mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)}) \approx \mu_N(\mathcal{E}_a^{(\mathfrak{p})})$ for all $a \in I$. By assumption (T3) for $k = \mathfrak{p}$, $\mu_N(y) \approx \mu_N(\mathcal{E}_a^{(\mathfrak{p})})$ for all $y \in \mathcal{E}_a^{(\mathfrak{p})}$, which proves the first claim of the lemma.

To prove the second claim, fix x, y in $\mathcal{E}_i^{(\mathfrak{p}+1)}$. If x, y belong to the same set $\mathcal{E}_a^{(\mathfrak{p})}$, the lemma follows from assumption (T4) for $k = \mathfrak{p}$, the first part of the lemma and the fact that $\theta_N(\mathfrak{p} - 1) \prec \theta_N(\mathfrak{p})$.

Assume that x, y belongs to different $\mathfrak{p}$-metastates, say $x \in \mathcal{E}_a^{(\mathfrak{p})}$, $y \in \mathcal{E}_b^{(\mathfrak{p})}$, $a \neq b$. Since $\mathcal{E}_a^{(\mathfrak{p})} \sim \mathcal{E}_b^{(\mathfrak{p})}$, there exists an open path, $a = a_0, \ldots, a_n = b$, from $\mathcal{E}_a^{(\mathfrak{p})}$ to $\mathcal{E}_b^{(\mathfrak{p})}$. This means that $\theta_N(\mathfrak{p}) r_N^{\mathcal{E}^{(\mathfrak{p})}}(\mathcal{E}_{a_m}^{(\mathfrak{p})}, \mathcal{E}_{a_{m+1}}^{(\mathfrak{p})})$ converges to a positive number for $0 \leq m < n$. Therefore, by (10.3), there exists a positive number $C_0 > 0$, independent of N and which may change from line to line, such that

$$\theta_N(\mathfrak{p}) \frac{\mathrm{cap}_N(\mathcal{E}_{a_m}^{(\mathfrak{p})}, \mathcal{E}_{a_{m+1}}^{(\mathfrak{p})})}{\mu_N(\mathcal{E}_{a_m}^{(\mathfrak{p})})} \geq C_0$$

for all N large enough and $0 \leq m < n$. Since by Lemma 4.1 $\mathrm{cap}_N(A, B) \approx G_N(A, B)$, $G_N(\mathcal{E}_{a_m}^{(\mathfrak{p})}, \mathcal{E}_{a_{m+1}}^{(\mathfrak{p})}) \geq C_0 \mu_N(\mathcal{E}_{a_m})/\theta_N(\mathfrak{p})$. There exists, in particular, a path γ_m from $x_m \in \mathcal{E}_{a_m}^{(\mathfrak{p})}$ to $y_{m+1} \in \mathcal{E}_{a_{m+1}}^{(\mathfrak{p})}$, $0 \leq m < n$, with $G_N(\gamma_m) \geq C_0 \mu_N(\mathcal{E}_{a_m}^{(\mathfrak{p})})/\theta_N(\mathfrak{p})$.

By assumption (T4) for $k = \mathfrak{p}$ and arguments similar to the ones used above, there exist: a path γ_0' from $x \in \mathcal{E}_a^{(\mathfrak{p})}$ to $x_0 \in \mathcal{E}_a^{(\mathfrak{p})}$ such that $G_N(\gamma_0') \geq C_0 \mu_N(\mathcal{E}_a^{(\mathfrak{p})})/\theta_N(\mathfrak{p} - 1)$; paths γ_m' from $y_m \in \mathcal{E}_{a_m}^{(\mathfrak{p})}$ to $x_m \in \mathcal{E}_{a_m}^{(\mathfrak{p})}$, $1 \leq m < n$, such that $G_N(\gamma_m') \geq C_0 \mu_N(\mathcal{E}_{a_m}^{(\mathfrak{p})})/\theta_N(\mathfrak{p} - 1)$; and a path γ_n' from $y_n \in \mathcal{E}_b^{(\mathfrak{p})}$ to $y \in \mathcal{E}_b^{(\mathfrak{p})}$ such that $G_N(\gamma_n') \geq C_0 \mu_N(\mathcal{E}_b^{(\mathfrak{p})})/\theta_N(\mathfrak{p} - 1)$.

Since, by the first part of the lemma, $\mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)}) \approx \mu_N(\mathcal{E}_c^{(\mathfrak{p})})$ for all $c \in I$, juxtaposing all these paths, we obtain a path γ from x to y such that $G_N(\gamma) \geq C_0 \, \mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})/\theta_N(\mathfrak{p})$. This shows that

$$\mathrm{cap}_N(x, y) \approx G_N(\{x\}, \{y\}) \succeq G_N(\gamma) \geq \frac{C_0 \, \mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})}{\theta_N(\mathfrak{p})},$$

which proves the lemma. $\square$

We next show that condition (T7) is in force on any time scale longer than $\theta_N(\mathfrak{p})$.

Lemma 7.5. *Let $\{\theta_N : N \geq 1\}$ be a sequence such that $\theta_N \succ \theta_N(\mathfrak{p})$. Then for every $1 \leq i \leq \nu(\mathfrak{p}+1)$, $x \in \mathcal{E}_i^{(\mathfrak{p}+1)}$ and $\delta > 0$,*

$$\lim_{N \to \infty} \max_{y \in \mathcal{E}_i^{(\mathfrak{p}+1)}} \mathbf{P}_y^N[T_x > \delta \, \theta_N] = 0.$$

Proof. Fix $1 \leq i \leq \nu(\mathfrak{p}+1)$, $x, y \in \mathcal{E}_i^{(\mathfrak{p}+1)}$ and $\delta > 0$. Denote by $\mathcal{E}_a^{(\mathfrak{p})}, \mathcal{E}_b^{(\mathfrak{p})} \subset \mathcal{E}_i^{(\mathfrak{p}+1)}$ the $\mathfrak{p}$-metastates which contain x, y, respectively. Since $\theta_N \succ \theta_N(\mathfrak{p})$, by the strong Markov property, for every $t > 0$ and for every N large enough,

$$\mathbf{P}_y^N[T_x > \delta\theta_N] \leq \mathbf{P}_y^N[T_{\mathcal{E}_a^{(\mathfrak{p})}} > t \, \theta_N(\mathfrak{p})] + \max_{z \in \mathcal{E}_a^{(\mathfrak{p})}} \mathbf{P}_z^N[T_x > \delta\theta_N/2]. \tag{7.4}$$

We claim that both expressions vanish as $N \uparrow \infty$ and then $t \uparrow \infty$. Denote by $T_{\mathcal{E}_a^{(\mathfrak{p})}}^{(\mathfrak{p})}$ the time of hitting of $\mathcal{E}_a^{(\mathfrak{p})}$ by the trace process $\eta_t^{N,\mathfrak{p}}$ defined just before (7.2). The first term on the right hand side of the previous formula is bounded above by

$$\mathbf{P}_y^N\left[\int_0^t \mathbf{1}\{\eta_{s\theta_N(\mathfrak{p})}^N \in \Delta_\mathfrak{p}\} \, ds > \epsilon\right] + \mathbf{P}_y^N\left[T_{\mathcal{E}_a^{(\mathfrak{p})}}^{(\mathfrak{p})} > (t - \epsilon) \, \theta_N(\mathfrak{p})\right]$$

for every $0 < \epsilon < t$. By property (T8) for $k = \mathfrak{p}$, the first term vanishes as $N \uparrow \infty$ for every $\epsilon > 0$. By the convergence of the process $\Psi_\mathfrak{p}(\eta_{t\theta_N(\mathfrak{p})}^{N,\mathfrak{p}})$ to the Markov process with rates $\mathfrak{r}_\mathfrak{p}(i, j)$, assumed in (T6), the second term converges as $N \uparrow \infty$ to $\mathbb{P}_b[T_a > (t - \epsilon)]$, where T_a stands for the hitting time of a. Since $\mathcal{E}_i^{(\mathfrak{p}+1)}$ is a leaf, the asymptotic dynamics is an irreducible Markov process on the set of indices $c \in \{1 \ldots, \nu(\mathfrak{p})\}$ such that $\mathcal{E}_c^{(\mathfrak{p})} \subset \mathcal{E}_i^{(\mathfrak{p}+1)}$. In particular, $\mathbb{P}_b[T_a > (t - \epsilon)]$ vanishes as $t \uparrow \infty$. This proves that the first term in (7.4) vanishes as $N \uparrow \infty$ and then $t \uparrow \infty$.

The second term in (7.4) vanishes as $N \uparrow \infty$ by property (T7) for $k = \mathfrak{p}$. This proves the lemma. $\square$

The next lemma shows that we may from now on restrict our attention to the trace process $\{\eta_t^{N,\mathfrak{p}} : t \geq 0\}$ in our investigation of the metastability of $\{\eta_t^N : t \geq 0\}$ on a time scale longer than $\theta_N(\mathfrak{p})$.

Lemma 7.6. *Assume that the trace processes $\{\eta_t^{N,\mathfrak{p}} : t \geq 0\}$ satisfy condition (T8) on some time scale $\theta_N \succ \theta_N(\mathfrak{p})$ and for some subset $\Delta_{\mathfrak{p}+1}^*$ of $\mathcal{E}^{(\mathfrak{p})}$:*

$$\lim_{N \to \infty} \max_{x \in \mathcal{E}^{(\mathfrak{p})}} \mathbf{E}_x^N\left[\int_0^t \mathbf{1}\{\eta_{s\theta_N}^{N,\mathfrak{p}} \in \Delta_{\mathfrak{p}+1}^*\} \, ds\right] = 0.$$

Then, the same property holds for the Markov process $\{\eta_t^N : t \geq 0\}$ with $\Delta_{\mathfrak{p}} \cup \Delta_{\mathfrak{p}+1}^*$ in place of $\Delta_{\mathfrak{p}+1}^*$:

$$\lim_{N \to \infty} \max_{x \in E_N} \mathbf{E}_x^N \left[\int_0^t \mathbf{1}\{\eta_{s\theta_N}^N \in \Delta_{\mathfrak{p}} \cup \Delta_{\mathfrak{p}+1}^*\} \, ds \right] = 0.$$

Proof. Fix $x \in E$ and observe that

$$\mathbf{E}_x^N \left[\int_0^t \mathbf{1}\{\eta_{s\theta_N}^N \in \Delta_{\mathfrak{p}} \cup \Delta_{\mathfrak{p}+1}^*\} \, ds \right] \leq \mathbf{E}_x^N \left[\int_0^t \mathbf{1}\{\eta_{s\theta_N}^N \in \Delta_{\mathfrak{p}}\} \, ds \right]$$

$$+ \max_{y \in \mathcal{E}^{(\mathfrak{p})}} \mathbf{E}_y^N \left[\int_0^t \mathbf{1}\{\eta_{s\theta_N}^{N,\mathfrak{p}} \in \Delta_{\mathfrak{p}+1}^*\} \, ds \right].$$

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

$$\frac{\theta_N(\mathfrak{p})}{\theta_N} \sum_{n=0}^{[\theta_N/\theta_N(\mathfrak{p})]} \mathbf{E}_x^N \left[\int_{nt}^{(n+1)t} \mathbf{1}\{\eta_{s\theta_N(\mathfrak{p})}^N \in \Delta_{\mathfrak{p}}\} \, ds \right],$$

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

$$2 \max_{y \in E} \mathbf{E}_y^N \left[\int_0^t \mathbf{1}\{\eta_{s\theta_N(\mathfrak{p})}^N \in \Delta_{\mathfrak{p}}\} \, ds \right],$$

which vanishes as $N \uparrow \infty$ in virtue of (T8) for $k = \mathfrak{p}$. $\square$

Consider the trace process $\{\eta_t^{N,\mathfrak{p}} : t \geq 0\}$. By formula (6.12) in [2], its invariant probability measure is the measure μ_N conditioned to $\mathcal{E}^{(\mathfrak{p})}$, and by [2, Lemma 6.9] the capacity between two disjoint subsets of $\mathcal{E}^{(\mathfrak{p})}$ for the trace process $\{\eta_t^{N,\mathfrak{p}} : t \geq 0\}$ is equal to the capacity for the original process divided by $\mu_N(\mathcal{E}^{(\mathfrak{p})})$.

The evolution the trace process $\{\eta_t^{N,\mathfrak{p}} : t \geq 0\}$ on $\mathcal{E}^{(\mathfrak{p})}$ is similar to that of $\{\eta_t^N : t \geq 0\}$ among the shallowest valleys. We claim, for instance, that $(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)}, x), x \in \mathcal{E}_i^{(\mathfrak{p}+1)}$, $1 \leq i \leq \nu(\mathfrak{p}+1)$, are valleys for the trace process $\{\eta_t^{N,\mathfrak{p}} : t \geq 0\}$. The proof of this assertion is divided into several steps. We first show that

$$G_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)}) \prec \frac{\mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})}{\theta_N(\mathfrak{p})}, \quad 1 \leq i \leq \nu(\mathfrak{p}+1). \tag{7.5}$$

Indeed, since $\mathcal{E}_i^{(\mathfrak{p}+1)}$ is a leaf, there is no open path from some $\mathcal{E}_i^{(\mathfrak{p})} \subset \mathcal{E}_i^{(\mathfrak{p}+1)}$ to some $\mathcal{E}_b^{(\mathfrak{p})} \not\subset \mathcal{E}_i^{(\mathfrak{p}+1)}$. Therefore, since by Lemma 7.4 $\mu_N(x) \approx \mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})$, $x \in \mathcal{E}_i^{(\mathfrak{p}+1)}$, by the definition of the average rate,

$$\lim_{N \to \infty} \theta_N(\mathfrak{p}) \, r_N^{\mathcal{E}^{(\mathfrak{p})}} \left(\mathcal{E}_i^{(\mathfrak{p}+1)}, \bigcup_b \mathcal{E}_b^{(\mathfrak{p})} \right) = \lim_{N \to \infty} \theta_N(\mathfrak{p}) \sum_a r_N^{\mathcal{E}^{(\mathfrak{p})}} \left(\mathcal{E}_a^{(\mathfrak{p})}, \bigcup_b \mathcal{E}_b^{(\mathfrak{p})} \right) = 0,$$

where the sum is taken over all $\mathfrak{p}$-metastates $\mathcal{E}_a^{(\mathfrak{p})} \subset \mathcal{E}_i^{(\mathfrak{p}+1)}$ and the union over all $\mathfrak{p}$-metastates $\mathcal{E}_b^{(\mathfrak{p})} \not\subset \mathcal{E}_i^{(\mathfrak{p}+1)}$. Hence, by [2, Lemma 6.7] and Lemma 4.1,

$$\lim_{N \to \infty} \theta_N(\mathfrak{p}) \frac{G_N\left(\mathcal{E}_i^{(\mathfrak{p}+1)}, \bigcup_b \mathcal{E}_b^{(\mathfrak{p})}\right)}{\mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})} = 0. \tag{7.6}$$

This proves (7.5) in view of (4.2) and because $\breve{\mathcal{E}}_i^{(\mathfrak{p}+1)} \subset \cup_b \mathcal{E}_b^{(\mathfrak{p})}$.

Recall the definition of the set $\Delta_{\mathfrak{p}+1}^o$ introduced just before Lemma 7.4. Denote by $\mathcal{B}_i^{(\mathfrak{p}+1)}$, $1 \leq i \leq \nu(\mathfrak{p}+1)$, the union of all $\mathfrak{p}$-metastates $\mathcal{E}_b^{(\mathfrak{p})}$ which have measure of lower magnitude than $\mathcal{E}_i^{(\mathfrak{p}+1)}$ and which are contained in $\Delta_{\mathfrak{p}+1}^o$. Let also

$$\mathcal{F}_i^{(\mathfrak{p}+1)} = \mathcal{E}^{(\mathfrak{p})} \setminus \left[\mathcal{E}_i^{(\mathfrak{p}+1)} \cup \mathcal{B}_i^{(\mathfrak{p}+1)}\right], \quad 1 \leq i \leq \nu(\mathfrak{p}+1).$$

Lemma 7.7. *Fix $1 \leq i \leq \nu(\mathfrak{p}+1)$ and x in $\mathcal{E}_i^{(\mathfrak{p}+1)}$. The triple $(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{E}_i^{(\mathfrak{p}+1)} \cup \mathcal{B}_i^{(\mathfrak{p}+1)}, x)$ is a valley for the trace process $\{\eta_t^{N,\mathfrak{p}} : t \geq 0\}$ of depth $\theta_{N,i} = \mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})/\mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)})$. Moreover, $\theta_{N,i} \succ \theta_N(\mathfrak{p})$.*

Proof. Fix $1 \leq i \leq \nu(\mathfrak{p}+1)$ and x in $\mathcal{E}_i^{(\mathfrak{p}+1)}$. In view of Theorem 2.6, formula (6.12) and Lemma 6.9 in [2], we only need to check that

$$\lim_{N \to \infty} \max_{y \in \mathcal{E}_i^{(\mathfrak{p}+1)}} \frac{\mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)})}{\mathrm{cap}_N(x, y)} = 0.$$

This follows from Lemmas 7.4 and 4.1, (4.2) and (7.6).

It remains to show that $\theta_{N,i} \succ \theta_N(\mathfrak{p})$. Since $\mathcal{F}_i^{(\mathfrak{p}+1)}$ is contained in $\cup_b \mathcal{E}_b^{(\mathfrak{p})}$, where the union is performed over all $\mathfrak{p}$-metastates which are not contained in $\mathcal{E}_i^{(\mathfrak{p}+1)}$, and since $\mathrm{cap}_N(A, B) \leq \mathrm{cap}_N(A, C)$ if $B \subset C$, by Lemma 4.1, $\theta_N(\mathfrak{p})/\theta_{N,i}$ is bounded above by

$$C_1 \theta_N(\mathfrak{p}) \frac{G_N\left(\mathcal{E}_i^{(\mathfrak{p}+1)}, \bigcup_b \mathcal{E}_b^{(\mathfrak{p})}\right)}{\mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})}$$

for some finite constant C_1 independent of N. By (7.6) this expression vanishes as $N \uparrow \infty$. $\square$

Denote by $\mathbf{P}_x^{N,\mathfrak{p}}$, $x \in \mathcal{E}^{(\mathfrak{p})}$, the probability on the path space $D(\mathbb{R}_+, \mathcal{E}^{(\mathfrak{p})})$ induced by the trace process $\{\eta_t^{N,\mathfrak{p}} : t \geq 0\}$ starting from x.

Lemma 7.8. *Fix $1 \leq i \leq \nu(\mathfrak{p}+1)$ and x in $\mathcal{E}_i^{(\mathfrak{p}+1)}$. The triple $(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{E}_i^{(\mathfrak{p}+1)} \cup \Delta_{\mathfrak{p}+1}^o, x)$ is a valley for the trace process $\{\eta_t^{N,\mathfrak{p}} : t \geq 0\}$ of depth $\theta_{N,i} = \mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})/\mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)})$. Moreover, $\mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)}) \approx \mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})$.*

Proof. Fix $1 \leq i \leq \nu(\mathfrak{p}+1)$ and recall the definition of $\theta_{N,i}$ introduced in Lemma 7.7. By this lemma and by Lemma 10.1, to prove the first assertion we need to show that for every $\delta > 0$,

$$\lim_{N \to \infty} \max_{y \in \Delta_{\mathfrak{p}+1}^o \setminus \mathcal{B}_i^{(\mathfrak{p}+1)}} \mathbf{P}_y^{N,\mathfrak{p}}\left[T_{\breve{\mathcal{E}}_i^{(\mathfrak{p}+1)}} > \delta \theta_{N,i}\right] = 0.$$

Since, by Lemma 7.7, $\theta_{N,i} \succ \theta_N(\mathfrak{p})$, it is enough to show that

$$\lim_{t\to\infty} \lim_{N\to\infty} \max_{y\in\Delta^o_{\mathfrak{p}+1}\setminus\mathcal{B}^{(\mathfrak{p}+1)}_i} \mathbf{P}^{N,\mathfrak{p}}_y\big[T_{\breve{\mathcal{E}}^{(\mathfrak{p}+1)}_i} > t\,\theta_N(\mathfrak{p})\big] = 0.$$

Fix $y \in \Delta^o_{\mathfrak{p}+1} \setminus \mathcal{B}^{(\mathfrak{p}+1)}_i$. By definition, y belongs to some $\mathfrak{p}$-metastate $\mathcal{E}^{(\mathfrak{p})}_b \not\subset \mathcal{E}^{(\mathfrak{p}+1)}_i$ and $\mu_N(\mathcal{E}^{(\mathfrak{p})}_b) \succeq \mu_N(\mathcal{E}^{(\mathfrak{p}+1)}_i)$. We claim that there is no open path from $\mathcal{E}^{(\mathfrak{p})}_b$ to $\mathcal{E}^{(\mathfrak{p}+1)}_i$.

Indeed, suppose that there is an open path. In this case, since $\mu_N(\mathcal{E}^{(\mathfrak{p})}_b) \succeq \mu_N(\mathcal{E}^{(\mathfrak{p}+1)}_i)$, by (7.1), we necessarily have $\mu_N(\mathcal{E}^{(\mathfrak{p})}_b) \approx \mu_N(\mathcal{E}^{(\mathfrak{p}+1)}_i)$. Considering the last two $\mathfrak{p}$-metastates of the open path from $\mathcal{E}^{(\mathfrak{p})}_b$ to $\mathcal{E}^{(\mathfrak{p}+1)}_i$, we find a $\mathfrak{p}$-metastate $\mathcal{E}^{(\mathfrak{p})}_c \not\subset \mathcal{E}^{(\mathfrak{p}+1)}_i$, $\mu_N(\mathcal{E}^{(\mathfrak{p})}_c) \approx \mu_N(\mathcal{E}^{(\mathfrak{p}+1)}_i)$, and a $\mathfrak{p}$-metastate $\mathcal{E}^{(\mathfrak{p})}_a \subset \mathcal{E}^{(\mathfrak{p}+1)}_i$ such that $\mathfrak{r}_\mathfrak{p}(c,a) > 0$. Therefore, by (7.1) and by reversibility,

$$\mathfrak{r}_\mathfrak{p}(a,c) = \lim_{N\to\infty} \theta_N(\mathfrak{p})\, r^{\mathcal{E}^{(\mathfrak{p})}}_N(\mathcal{E}^{(\mathfrak{p})}_a, \mathcal{E}^{(\mathfrak{p})}_c)$$

$$= \lim_{N\to\infty} \frac{\mu_N(\mathcal{E}^{(\mathfrak{p})}_c)}{\mu_N(\mathcal{E}^{(\mathfrak{p})}_a)} \theta_N(\mathfrak{p})\, r^{\mathcal{E}^{(\mathfrak{p})}}_N(\mathcal{E}^{(\mathfrak{p})}_c, \mathcal{E}^{(\mathfrak{p})}_a) = \mathfrak{r}_\mathfrak{p}(a,c) \lim_{N\to\infty} \frac{\mu_N(\mathcal{E}^{(\mathfrak{p})}_c)}{\mu_N(\mathcal{E}^{(\mathfrak{p})}_a)} > 0,$$

which contradicts the fact that $\mathcal{E}^{(\mathfrak{p}+1)}_i$ is a leaf.

By (7.2) with $k = \mathfrak{p}$, starting from y the process $X^{N,\mathfrak{p}}_{t\theta_N(\mathfrak{p})}$ converges to the Markov process on $\{1, \ldots, \nu(\mathfrak{p})\}$ with rates $\mathfrak{r}_\mathfrak{p}$ starting from b. Therefore,

$$\lim_{N\to\infty} \mathbf{P}^{N,\mathfrak{p}}_y\big[T_{\breve{\mathcal{E}}^{(\mathfrak{p}+1)}_i} > t\,\theta_N(\mathfrak{p})\big] \le \mathbb{P}_b[T_A > t],$$

where $A = \{c : \mathcal{E}^{(\mathfrak{p})}_c \subset \breve{\mathcal{E}}^{(\mathfrak{p}+1)}_i\}$. Since there is no open path from $\mathcal{E}^{(\mathfrak{p})}_b$ to $\mathcal{E}^{(\mathfrak{p}+1)}_i$ and since $\mathcal{E}^{(\mathfrak{p})}_b \subset \Delta^o_{\mathfrak{p}+1}$, the state b is transient for the Markov process on $\{1, \ldots, \nu(\mathfrak{p})\}$ with rates $\mathfrak{r}_\mathfrak{p}$ and all its limit points are contained in A. Hence, $\mathbb{P}_b[T_A > t]$ vanishes as $t \uparrow \infty$. This proves the first assertion of the lemma.

To prove the second statement, note that $\mathrm{cap}_N(\mathcal{E}^{(\mathfrak{p}+1)}_i, \mathcal{F}^{(\mathfrak{p}+1)}_i) \succeq \mathrm{cap}_N(\mathcal{E}^{(\mathfrak{p}+1)}_i, \breve{\mathcal{E}}^{(\mathfrak{p}+1)}_i)$ because $\breve{\mathcal{E}}^{(\mathfrak{p}+1)}_i \subset \mathcal{F}^{(\mathfrak{p}+1)}_i$.

By Lemma 4.1, to prove the reverse inequality we may replace the capacities by the function G_N. There exists a path $\gamma = (x_0, \ldots, x_n)$ from $\mathcal{E}^{(\mathfrak{p}+1)}_i$ to $\mathcal{F}^{(\mathfrak{p}+1)}_i$ such that $G_N(\gamma) = G_N(\mathcal{E}^{(\mathfrak{p}+1)}_i, \mathcal{F}^{(\mathfrak{p}+1)}_i)$. If x_n belongs to $\breve{\mathcal{E}}^{(\mathfrak{p}+1)}_i$, we have that $G_N(\gamma) \le G_N(\mathcal{E}^{(\mathfrak{p}+1)}_i, \breve{\mathcal{E}}^{(\mathfrak{p}+1)}_i)$ and the statement is proved.

If, on the other hand, x_n belongs to some metastate $\mathcal{E}^{(\mathfrak{p})}_b \subset \Delta^o_{\mathfrak{p}+1} \setminus \mathcal{B}^{(\mathfrak{p}+1)}_i$ we proceed as follows. We have already shown in the first part of the proof that there exists an open path from $\mathcal{E}^{(\mathfrak{p})}_b$ to $\breve{\mathcal{E}}^{(\mathfrak{p}+1)}_i$. Repeating the arguments presented in the proof of Lemma 7.4 and keeping in mind the second assertion of (7.1), we show that there exists a path $\tilde{\gamma}$ from x_n to $\breve{\mathcal{E}}^{(\mathfrak{p}+1)}_i$ such that $G_N(\tilde{\gamma}) \ge C_0\mu_N(x_n)/\theta_N(\mathfrak{p})$ for some finite constant C_0 independent of N. By the definition of $\mathcal{B}^{(\mathfrak{p}+1)}_i$, this latter expression is bounded below by $C_0\mu_N(\mathcal{E}^{(\mathfrak{p}+1)}_i)/\theta_N(\mathfrak{p})$. By (7.6), $G_N(\gamma) = G_N(\mathcal{E}^{(\mathfrak{p}+1)}_i, \mathcal{F}^{(\mathfrak{p}+1)}_i) \prec \mu_N(\mathcal{E}^{(\mathfrak{p}+1)}_i)/\theta_N(\mathfrak{p})$. Hence, if we denote by $\gamma \oplus \tilde{\gamma}$ the juxtaposition of γ and $\tilde{\gamma}$, we have a path $\gamma \oplus \tilde{\gamma}$ from $\mathcal{E}^{(\mathfrak{p}+1)}_i$ to $\breve{\mathcal{E}}^{(\mathfrak{p}+1)}_i$ such that $G_N(\gamma \oplus \tilde{\gamma}) = G_N(\gamma) = G_N(\mathcal{E}^{(\mathfrak{p}+1)}_i, \mathcal{F}^{(\mathfrak{p}+1)}_i)$. This proves the second assertion of the lemma. $\square$

It follows from the two previous lemmas that the depth $\theta_{N,i}$ of the valley $(\mathcal{E}^{(\mathfrak{p}+1)}_i, \mathcal{E}^{(\mathfrak{p}+1)}_i \cup \Delta^o_{\mathfrak{p}+1}, x)$, $1 \le i \le \nu(\mathfrak{p}+1)$, is of the same magnitude as $\mu_N(\mathcal{E}^{(\mathfrak{p}+1)}_i)/\mathrm{cap}_N(\mathcal{E}^{(\mathfrak{p}+1)}_i, \breve{\mathcal{E}}^{(\mathfrak{p}+1)}_i)$ and much larger than $\theta_N(\mathfrak{p})$.

Fix a subset I of $\{1, \ldots, \nu(\mathfrak{p} + 1)\}$ and let $J = \{1, \ldots, \nu(\mathfrak{p} + 1)\} \setminus I$, $\mathcal{E}_{K,\mathfrak{p}+1} = \cup_{i \in K} \mathcal{E}_i^{(\mathfrak{p}+1)}$, $K \subset \{1, \ldots, \nu(\mathfrak{p} + 1)\}$. By Lemma 4.3, the following limit exists:

$$f_{I,J}^{(\mathfrak{p}+1)}(x) := \lim_{N \to \infty} \mathbf{P}_x^N[T_{\mathcal{E}_{I,\mathfrak{p}+1}} < T_{\mathcal{E}_{J,\mathfrak{p}+1}}].$$

In particular, by Lemma 4.2,

$$\lim_{N \to \infty} \frac{\mathrm{cap}_N(\mathcal{E}_I^{(\mathfrak{p}+1)}, \mathcal{E}_J^{(\mathfrak{p}+1)})}{\mathfrak{g}_N(\mathcal{E}_I^{(\mathfrak{p}+1)}, \mathcal{E}_J^{(\mathfrak{p}+1)})} = \frac{1}{2} \sum g(x, y) \, [f_{I,J}^{(\mathfrak{p}+1)}(y) - f_{I,J}^{(\mathfrak{p}+1)}(x)]^2 \in (0, \infty),$$

where the sum on the right hand side is taken over all pairs $(x, y) \in \mathfrak{B}(\mathcal{E}_I^{(\mathfrak{p}+1)}, \mathcal{E}_J^{(\mathfrak{p}+1)})$.

For the same reasons, the limit

$$f_i^{(\mathfrak{p}+1)}(x) := \lim_{N \to \infty} \mathbf{P}_x^N\left[T_{\mathcal{E}_i^{(\mathfrak{p}+1)}} < T_{\mathcal{F}_i^{(\mathfrak{p}+1)}}\right]$$

exists and

$$\lim_{N \to \infty} \frac{\mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)})}{\mathfrak{g}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)})} = \frac{1}{2} \sum g(x, y) \, [f_i^{(\mathfrak{p}+1)}(y) - f_i^{(\mathfrak{p}+1)}(x)]^2 \in (0, \infty),$$

where the sum on the right hand side is taken over all pairs $(x, y) \in \mathfrak{B}(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)})$.

Let $g_i^{(\mathfrak{p}+1)} = f_{I,J}^{(\mathfrak{p}+1)}$ for $I = \{i\}$. We claim that $g_i^{(\mathfrak{p}+1)} = f_i^{(\mathfrak{p}+1)}$, in other words, that for all $x \in E$,

$$\lim_{N \to \infty} \mathbf{P}_x^N[T_{\mathcal{E}_i^{(\mathfrak{p}+1)}} < T_{\mathcal{F}_i^{(\mathfrak{p}+1)}}] = \lim_{N \to \infty} \mathbf{P}_x^N[T_{\mathcal{E}_i^{(\mathfrak{p}+1)}} < T_{\breve{\mathcal{E}}_i^{(\mathfrak{p}+1)}}]. \tag{7.7}$$

Indeed, fix $x \in E$. Since $\lim_{N \to \infty} \mathbf{P}_x^N[T_{\mathcal{E}^{(\mathfrak{p})}} = T_y]$, $y \in \mathcal{E}^{(\mathfrak{p})}$, exists by Lemma 4.3, and since all sets involved are contained in $\mathcal{E}^{(\mathfrak{p})}$, taking conditional expectation with respect to $T_{\mathcal{E}^{(\mathfrak{p})}}$ and applying the strong Markov property, to prove (7.7) it is enough to show that for all $y \in \mathcal{E}^{(\mathfrak{p})}$,

$$\lim_{N \to \infty} \mathbf{P}_y^N[T_{\mathcal{E}_i^{(\mathfrak{p}+1)}} < T_{\mathcal{F}_i^{(\mathfrak{p}+1)}}] = \lim_{N \to \infty} \mathbf{P}_y^N[T_{\mathcal{E}_i^{(\mathfrak{p}+1)}} < T_{\breve{\mathcal{E}}_i^{(\mathfrak{p}+1)}}].$$

At this point we may replace the process η_t^N by the trace process $\eta_t^{N,\mathfrak{p}}$. Since $\breve{\mathcal{E}}_i^{(\mathfrak{p}+1)}$ is contained in $\mathcal{F}_i^{(\mathfrak{p}+1)}$, by the strong Markov property, to prove the previous identity we have to show that for every $z \in \mathcal{F}_i^{(\mathfrak{p}+1)}$,

$$\lim_{N \to \infty} \mathbf{P}_z^N[T_{\mathcal{E}_i^{(\mathfrak{p}+1)}} < T_{\breve{\mathcal{E}}_i^{(\mathfrak{p}+1)}}] = 0.$$

Since there is no open path from $\mathcal{F}_i^{(\mathfrak{p}+1)}$ to $\mathcal{E}_i^{(\mathfrak{p}+1)}$, since by (7.2) the speeded-up blind process $X_t^{N,\mathfrak{p}}$ converges to the Markov process with rates $\mathfrak{r}_\mathfrak{p}$ whose recurrent states are the indices $a \in \{1, \ldots, \nu(\mathfrak{p})\}$ such that $\mathcal{E}_a^{(\mathfrak{p})} \subset \mathcal{E}^{(\mathfrak{p}+1)}$, the previous identity holds, proving claim (7.7).

We proved in Lemma 7.8 that $\mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)}) \approx \mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})$. Hence, by Lemma 4.1, $G_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)}) \approx G_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})$. In particular, $\mathfrak{g}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)}) = \mathfrak{g}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})$ and, in consequence, $\mathfrak{B}(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)}) = \mathfrak{B}(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})$.

If follows from the previous considerations that

$$\lim_{N\to\infty} \frac{\mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})}{\mathfrak{g}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})} = \lim_{N\to\infty} \frac{\mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)})}{\mathfrak{g}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})} \in (0, \infty),$$

so

$$\lim_{N\to\infty} \frac{\mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})}{\mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{F}_i^{(\mathfrak{p}+1)})} = 1.$$

In consequence, by Lemma 7.8, the following result holds.

Lemma 7.9. *Fix $1 \le i \le \nu(\mathfrak{p}+1)$ and x in $\mathcal{E}_i^{(\mathfrak{p}+1)}$. The triple $(\mathcal{E}_i^{(\mathfrak{p}+1)}, \mathcal{E}_i^{(\mathfrak{p}+1)} \cup \Delta_{\mathfrak{p}+1}^o, x)$ is a valley for the trace process $\{\eta_t^{N,\mathfrak{p}} : t \ge 0\}$ of depth $\theta_{N,i} = \mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})/\mathrm{cap}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})$. Moreover,*

$$u_{\mathfrak{p}+1,i} := \lim_{N\to\infty} \frac{\mathfrak{g}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})}{\mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})} \theta_{N,i} \in (0, \infty).$$

Since the sequences $\mathfrak{g}_N(\mathcal{E}_i^{(\mathfrak{p}+1)}, \breve{\mathcal{E}}_i^{(\mathfrak{p}+1)})/\mu_N(\mathcal{E}_i^{(\mathfrak{p}+1)})$, $1 \le i \le \nu(\mathfrak{p}+1)$, are comparable, repeating the arguments presented in the proof of Proposition 5.8 we deduce the next result.

Lemma 7.10. *The sequences $\{\theta_{N,i} : N \ge 1\}$, $1 \le i \le \nu(\mathfrak{p}+1)$, are comparable.*

Let $\theta_N(\mathfrak{p}+1) = \min\{\theta_{N,i} : 1 \le i \le \nu(\mathfrak{p}+1)\}$ and let $S_{\mathfrak{p}+1} = \{i : \theta_{N,i} \approx \theta_N(\mathfrak{p}+1)\}$. Observe that $\theta_N(\mathfrak{p}) \prec \theta_N(\mathfrak{p}+1)$ and that (T5) holds for $k = \mathfrak{p}+1$ with this definition.

Denote by $X_t^{N,\mathfrak{p}+1} = \Psi_k(\eta_{t\theta_N(\mathfrak{p}+1)}^{N,\mathfrak{p}+1})$ the speeded-up blind process introduced in the statement of Theorem 2.1.

Lemma 7.11. *Condition* (T6) *holds for $k = \mathfrak{p}+1$.*

Proof. The arguments presented in Section 6 until Lemma 6.1 apply to the present context and show that conditions (7.1) are fulfilled for $k = \mathfrak{p}+1$.

It remains to prove the convergence of $X_t^{N,\mathfrak{p}+1}$. We need to check that the assumptions of [2, Theorem 2.7] are fulfilled. On the one hand, condition (H1) follows from condition (T4) for $k = \mathfrak{p}+1$, proved in Lemma 7.4, and from the fact that $\theta_{N,i} \succeq \theta_N(\mathfrak{p}+1) \succ \theta_N(\mathfrak{p})$, proved right after Lemma 7.10. On the other hand, condition (H0) is part of (7.1) which has already been proven. $\square$

To conclude the recurrence argument, it remains to show that property (T8) holds for $k = \mathfrak{p}+1$. We first show that it holds for the trace process $\eta_t^{N,\mathfrak{p}}$.

Lemma 7.12. *For all $t > 0$,*

$$\lim_{N\to\infty} \max_{x\in\mathcal{E}^{(\mathfrak{p})}} \mathbf{E}_x^N \left[\int_0^t \mathbf{1}\{\eta_{s\theta_N(\mathfrak{p}+1)}^{N,\mathfrak{p}} \in \Delta_{\mathfrak{p}+1}^o\} \, ds \right] = 0.$$

Proof. Since $\theta_N(\mathfrak{p}) \prec \theta_N(\mathfrak{p}+1)$, a change of variables in the time integral and the Markov property show that for every $T > 0$ and for every N large enough,

$$\mathbf{E}_x^N \left[\int_0^t \mathbf{1}\{\eta_{s\theta_N(\mathfrak{p}+1)}^{N,\mathfrak{p}} \in \Delta_{\mathfrak{p}+1}^o\} \, ds \right] \le \frac{2t}{T} \max_{y\in\mathcal{E}^{(\mathfrak{p})}} \mathbf{E}_y^N \left[\int_0^T \mathbf{1}\{\eta_{s\theta_N(\mathfrak{p})}^{N,\mathfrak{p}} \in \Delta_{\mathfrak{p}+1}^o\} \, ds \right]$$

for every $x \in \mathcal{E}^{(\mathfrak{p})}$. Note that the process on the right hand side is speeded up by $\theta_N(\mathfrak{p})$ and no longer by $\theta_N(\mathfrak{p}+1)$.

We estimate the expression on the right hand side of the previous formula. We may, of course, restrict the maximum to $\Delta^o_{\mathfrak{p}+1}$. Let T_1 be the first time that the trace process hits $\mathcal{E}^{(\mathfrak{p}+1)}$ and let T_2 be the time that it takes for the process to return to $\Delta^o_{\mathfrak{p}+1}$ after T_1:

$$T_1 = T_{\mathcal{E}^{(\mathfrak{p}+1)}}, \qquad T_2 = \inf\{s > 0 : \eta^{N,\mathfrak{p}}_{T_1+s} \in \Delta^o_{\mathfrak{p}+1}\}.$$

Fix $x \in \Delta^o_{\mathfrak{p}+1}$ and note that

$$\mathbf{E}^N_x\left[\frac{1}{T}\int_0^T \mathbf{1}\{\eta^{N,\mathfrak{p}}_{s\theta_N(\mathfrak{p})} \in \Delta^o_{\mathfrak{p}+1}\}\,\mathrm{d}s\right] \leq \mathbf{P}^{N,\mathfrak{p}}_x[T_1 > t_0\theta_N(\mathfrak{p})] + \mathbf{P}^{N,\mathfrak{p}}_x[T_2 \leq T\theta_N(\mathfrak{p})] + \frac{t_0}{T}$$

for all $t_0 > 0$. We have proved, in Lemma 7.8 for instance, that the first term on the right hand side vanishes as $N \uparrow \infty$ and then $t_0 \uparrow \infty$. By the strong Markov property, the second term is bounded by $\max_{y \in \mathcal{E}^{(\mathfrak{p}+1)}} \mathbf{P}^{N,\mathfrak{p}}_y[T_{\Delta^o_{\mathfrak{p}+1}} \leq T\theta_N(\mathfrak{p})]$. Since there is no open path from $\mathcal{E}^{(\mathfrak{p}+1)}$ to $\Delta^o_{\mathfrak{p}+1}$ this probability vanishes as $N \uparrow \infty$ for all $T > 0$. This concludes the proof. $\square$

The next result follows from Lemmas 7.6 and 7.12 and concludes the proof of Theorem 7.1.

Corollary 7.13. *Condition* (T8) *holds for* $k = \mathfrak{p} + 1$:

$$\lim_{N\to\infty} \max_{x\in E} \mathbf{E}^N_x\left[\int_0^t \mathbf{1}\{\eta^N_{s\theta_N(\mathfrak{p}+1)} \in \Delta_{\mathfrak{p}+1}\}\,\mathrm{d}s\right] = 0.$$

We conclude this section with a remark. Fix a level $\mathfrak{q}$ and denote by $P_N(x, i, j)$, $1 \leq i \neq j \leq v(\mathfrak{q})$, $x \in \mathcal{E}^{(\mathfrak{q})}_i$, the hitting probabilities

$$P_N(x, i, j) := \mathbf{P}^N_x[T_{\mathcal{E}^{(\mathfrak{q})}_j} = T_{\breve{\mathcal{E}}^{(\mathfrak{q})}_i}]. \tag{7.8}$$

Recall from Lemma 7.9 that $\theta_{N,i} = \mu_N(\mathcal{E}^{(\mathfrak{q})}_i)/\mathrm{cap}_N(\mathcal{E}^{(\mathfrak{q})}_i, \breve{\mathcal{E}}^{(\mathfrak{q})}_i)$. It follows from Lemma 7.10 with $\mathfrak{q} = \mathfrak{p} + 1$ that $\theta_N(\mathfrak{q})/\theta_{N,i}$ converges to some number denoted by $\Lambda(i) \in [0, \infty)$. On the other hand, by Lemma 4.3, $P_N(x, i, j)$ converges to some $P(x, i, j) \in [0, 1]$. Since by Lemma 7.9 $(\mathcal{E}^{(\mathfrak{q})}_i, \mathcal{E}^{(\mathfrak{q})}_i \cup \Delta^o_{\mathfrak{q}}, y)$, $y \in \mathcal{E}^{(\mathfrak{q})}_i$, is a valley for the trace process $\eta^{N,\mathfrak{q}}_t$, it is not difficult to show that the limit $P(x, i, j)$ does not depend on the starting point x. Therefore, by Lemma 10.2, for any $1 \leq i \neq j \leq v(\mathfrak{q})$,

$$\mathfrak{r}_{\mathfrak{q}}(i, j) = \lim_{N\to\infty} \frac{\theta_N(\mathfrak{q})}{\theta_{N,i}} \lim_{N\to\infty} \mathbf{P}^N_x[T_{\mathcal{E}^{(\mathfrak{q})}_j} = T_{\breve{\mathcal{E}}^{(\mathfrak{q})}_i}]. \tag{7.9}$$

8. Valleys and hitting times of the Ising model at low temperature

The proof of Theorem 3.1 follows the strategy presented in the previous sections. As we have seen, the approach relies on the characterization of the shallowest valleys of the model and on the computation of the depths and the hitting times of these valleys. We present in this section the shallowest valleys of the Ising model at low temperature and some estimates of the capacities and the hitting times.

In the present context, a path $\gamma = (\eta_0, \ldots, \eta_p)$ is a sequence of configuration in Ω such that for each $0 \leq j < p$, $\eta_{j+1} = \eta_j^x$ for some $x \in \Lambda_L$. We shall say that two configurations ξ and η in Ω are neighbors if $\xi = \eta^x$ for some $x \in \Lambda_L$.

Lemma 8.1. *Fix a configuration σ in Ω_o, $\sigma \neq +\mathbf{1}, -\mathbf{1}$. For all $\beta > 0$,*

$$G_\beta(\{\sigma\}, \Omega_\sigma) = \begin{cases} \mu_\beta(\sigma) \, \mathrm{e}^{-\beta[\ell(\sigma)-1]h} & \text{if } \ell(\sigma) \leq n_0, \\ \mu_\beta(\sigma) \, \mathrm{e}^{-\beta(2-h)} & \text{otherwise.} \end{cases} \tag{8.1}$$

Moreover,

$$G_\beta(\{-\mathbf{1}\}, \Omega_{-\mathbf{1}}) = \mu_\beta(-\mathbf{1}) \, \mathrm{e}^{-\beta(8-3h)}, \qquad G_\beta(\{+\mathbf{1}\}, \Omega_{+\mathbf{1}}) = \mu_\beta(+\mathbf{1}) \, \mathrm{e}^{-\beta(8+3h)}.$$

Proof. Fix a configuration σ satisfying the assumptions of the lemma and assume that $\ell := \ell(\sigma) \leq n_0$. Fix a positive rectangle R of σ of size $\ell \times m$ and assume that $m \geq 3$. Consider the sequence of configurations $\sigma = \eta_0, \ldots, \eta_\ell$ obtained by first flipping the spin at a corner of the rectangle R and then flipping contiguous spins along the smaller side. The last configuration η_ℓ is the configuration σ where the rectangle R has been replaced by a rectangle $R' \subset R$ of size $\ell \times (m-1)$.

The configuration η_ℓ belongs to Ω_σ and the path γ to $\Gamma_{\{\sigma\}, \Omega_\sigma}$. A simple computation shows that $\mu_\beta(\eta_{\ell-1}) = \min\{\mu_\beta(\eta_k) : 0 \leq k \leq \ell\}$, so $G_\beta(\{\sigma\}, \Omega_\sigma) \geq G_\beta(\gamma) = \mu_\beta(\eta_{\ell-1}) = \mu_\beta(\sigma) \mathrm{e}^{-\beta(\ell-1)h}$.

To prove the reverse inequality, note that the configuration σ has five kinds of different neighbors σ^x. A simple computation shows that $\mu_\beta(\sigma^x) < \mu_\beta(\sigma) \mathrm{e}^{-\beta(\ell-1)h}$ in four cases because $\ell \leq n_0 < 2/h$. The only type where this inequality does not hold occurs when we flip the spin at a corner of a positive rectangle of σ.

To compute $G_\beta(\{\sigma\}, \Omega_\sigma)$ we need to maximize $G_\beta(\gamma)$ over all paths γ from σ to Ω_σ. The previous observations show that the unique possible paths are those where we start flipping the corner of a positive rectangle.

This argument can be iterated. At each step we are only allowed to flip a positive spin which has two negative neighbors. After k flips we reach configurations of measure $\mu_\beta(\sigma)\mathrm{e}^{-\beta k h}$. Since we are not allowed to pass the level $\mu_\beta(\sigma)\mathrm{e}^{-\beta(\ell-1)h}$, the only configurations in Ω_o which can be reached after ℓ flips are the ones where a rectangle R of length $\ell \times m$ is replaced by a rectangle $R' \subset R$ of length $\ell \times (m-1)$.

The case of a rectangle R of size 2×2 is treated in a similar way. In this case, once one corner is removed, the next spins of the square flip at rate 1 to reach the configuration where the square R is removed. This proves the lemma in the case $\ell(\sigma) \leq n_0$.

Assume now that $\ell(\sigma) > n_0$. Consider the path $\gamma = (\sigma = \eta_0, \ldots, \eta_m)$, where η_1 is the configuration obtained from σ by flipping a negative spin contiguous to a positive rectangle, and where η_{j+1} is obtained from η_j, $2 \leq j < m$, by flipping a negative spin surrounded by two positive spins. The final configuration η_m is reached when no negative spin has two positive neighbors. Clearly, $\mu_\beta(\eta_1) = \min\{\mu_\beta(\eta_k) : 0 \leq k \leq m\}$, so $G_\beta(\{\sigma\}, \Omega_\sigma) \geq G_\beta(\gamma) = \mu_\beta(\eta_1) = \mu_\beta(\sigma)\mathrm{e}^{-\beta(2-h)}$.

An argument similar to the one presented in the first part of the proof of this lemma shows that the path proposed is the optimal one. This concludes the proof of the first part of the lemma.

Consider the path $\gamma = (\sigma_0 = -\mathbf{1}, \sigma_1, \ldots, \sigma_4)$ where σ_{j+1} is the configuration obtained from σ_j, $0 \leq j \leq 3$, by flipping a negative spin from a site with the largest possible number of neighbors with a positive spin. Hence, $\sigma_4 \in \Omega_{-\mathbf{1}}$ is obtained from $-\mathbf{1}$ by flipping the

spins of a 2×2 square and $G_\beta(\gamma) = \mu_\beta(-\mathbf{1}) \, e^{-\beta(8-3h)}$. In particular, $G_\beta(\{-\mathbf{1}\}, \Omega_{-\mathbf{1}}) \geq \mu_\beta(-\mathbf{1}) \, e^{-\beta(8-3h)}$.

To prove the reverse inequality, consider a path $\gamma = (\sigma_0, \ldots, \sigma_p)$ from $-\mathbf{1}$ to $\Omega_{-\mathbf{1}}$. Let σ_i be the first configuration in the path γ which has three positive spins. A simple computation shows that $\mu_\beta(\sigma_i) \leq \mu_\beta(-\mathbf{1}) \, e^{-\beta(8-3h)}$. This proves that $G_\beta(\gamma) \leq \mu_\beta(-\mathbf{1}) \, e^{-\beta(8-3h)}$, so $G_\beta(\{-\mathbf{1}\}, \Omega_{-\mathbf{1}}) \leq \mu_\beta(-\mathbf{1}) \, e^{-\beta(8-3h)}$, which proves the penultimate assertion of the lemma. The last statement is proved in a similar way. $\quad\square$

Recall the definition of the transition probabilities $p(\sigma, \sigma')$, $\sigma \in \Omega_o$, $\sigma' \in \mathbb{S}(\sigma)$, introduced in (3.1), (3.3). For $\ell(\sigma) = 2$ and $\ell(\sigma) > n_0$, cases where $\mathbb{S}(\sigma) = \mathbb{D}(\sigma)$, let $q(\sigma, \sigma') = p(\sigma, \sigma')$. For $\sigma \in \Omega_o$, $3 \leq \ell(\sigma) \leq n_0$, $\sigma' \in \mathbb{D}(\sigma)$, let $q(\sigma, \sigma')$ be defined by

$$q(\sigma, \sigma') = \frac{1}{|\mathbb{D}(\sigma)|}.$$

Note that $q(\sigma, \sigma') = p(\sigma, \sigma')$ for $\sigma' \in \mathbb{D}(\sigma) \cap \mathbb{S}(\sigma) = \mathbb{D}(\sigma) \cap \Omega_{o, \ell(\sigma)-1}$.

Lemma 8.2. *Fix a configuration σ in Ω_o, $\sigma \neq +\mathbf{1}, -\mathbf{1}$, and a configuration $\sigma' \in \mathbb{D}(\sigma)$. Then,*

$$\lim_{\beta \to \infty} \mathbf{P}_\sigma^\beta[T_{\sigma'} = T_{\Omega_\sigma}] = q(\sigma, \sigma').$$

Proof. Fix a configuration σ satisfying the assumptions of the lemma and a configuration $\sigma' \in \mathbb{D}(\sigma)$ and assume that $3 \leq \ell(\sigma) \leq n_0$. Denote by $\mathbb{W}(\sigma, \sigma')$ the set of configurations in $\mathbb{W}(\sigma)$ which are equal to σ' when we flip the positive spin surrounded by three negative spins. Note that $|\mathbb{W}(\sigma, \sigma')| = \ell(\sigma)$.

We present the proof for $\ell(\sigma) = 3$, the other cases being analogous. Since $3 = \ell(\sigma) \leq n_0 < 2/h$, we have that $h < 2/3$. For a configuration η for which all positive spins are surrounded by at most two negative spins, let $F_1(\eta)$ be the set of all configurations obtained from η by flipping a positive spin surrounded by two negative spins.

Let $f_\beta(\eta) = \mathbf{P}_\eta^\beta[T_{\sigma'} = T_{\Omega_\sigma}]$ and denote by f a limit point of the sequence f_β, as $\beta \uparrow \infty$. We need to show that $f(\sigma) = 1/|\mathbb{D}(\sigma)|$. Since f_β is harmonic, a simple computation shows that

$$f_\beta(\sigma) = \frac{1}{|F_1(\sigma)|} \sum_{\xi \in F_1(\sigma)} f_\beta(\xi) + o_1(\beta), \tag{8.2}$$

where $o_1(\beta)$ is an expression absolutely bounded by $C_0 \exp\{-2\beta[1-h]\}$ for some finite constant C_0 independent of β which may change from line to line. It follows from this identity that $f(\sigma) = |F_1(\sigma)|^{-1} \sum_{\xi \in F_1(\sigma)} f(\xi)$.

A similar argument shows that $f(\eta) = f(\sigma)$ for any configuration η in $F_1(\sigma)$. Let $F_2(\sigma)$ be the set of configurations obtained from a configuration in $F_1(\sigma)$ by flipping a positive spin surrounded by two negative spins. For the same reasons, $f(\xi) = f(\sigma)$ for any configuration ξ in $F_2(\sigma) \setminus \mathbb{W}(\sigma)$. Fix now a configuration η in $\mathbb{W}(\sigma, \sigma')$. If η differs from σ' by a spin in a corner of a positive rectangle of σ, $f(\eta) = (1/2)[1 + f(\sigma)]$, while if η differs from σ' by a spin not in a corner, $f(\eta) = (1/3)[1 + 2f(\sigma)]$. For a configuration η in $\mathbb{W}(\sigma) \setminus \mathbb{W}(\sigma, \sigma')$, if η differs from σ' by a spin in a corner of a positive rectangle of σ, $f(\eta) = (1/2)f(\sigma)$, while if η differs from σ' by a spin not in a corner, $f(\eta) = (2/3)f(\sigma)$.

Finally, observe that on applying the harmonic identity to the terms $f_\beta(\xi)$ in Eq. (8.2), after some elementary computations we obtain that

$$\sum_{\xi \in F_1(\sigma)} \sum_{\eta \in F_1(\xi)} \{f_\beta(\eta) - f_\beta(\sigma)\} = o_2(\beta),$$

where $o_2(\beta)$ is absolutely bounded by $C_0\{e^{-\beta h} + e^{-\beta[2-3h]}\}$. Since $h < 2/3$, the right hand side vanishes as $\beta \uparrow \infty$, so $\sum_{\eta \in F_2(\sigma)} \{f(\eta) - f(\sigma)\} = 0$. By the previous identities, this relation is reduced to $\sum_{\eta \in \mathbb{W}(\sigma)} \{f(\eta) - f(\sigma)\} = 0$. From this identity and the explicit values of f in $\mathbb{W}(\sigma)$, we obtain that $f(\sigma) = 1/|\mathbb{D}(\sigma)|$, which proves the lemma.

Suppose now that $\ell(\sigma) = 2 \leq n_0$ and note that Eq. (8.2) holds. The argument is analogous to the previous one, with one difference. If $\xi \in F_1(\sigma)$ is a configuration in which a spin of a 2×2 positive square Q of σ has been flipped, we have that $3f(\xi) = f(\sigma) + f(\eta_1) + f(\eta_2)$, where f is any limit point of the sequence f_β and η_1, η_2 are configurations obtained from σ by flipping a row or a column of the square Q. Iterating the argument on the basis of the harmonicity of f_β, we conclude that $3f(\xi) = f(\sigma) + 2f(\sigma^*)$, where σ^* is the configuration obtained from σ by flipping all spins of Q.

The proof for $\ell(\sigma) > n_0$ is similar. Observe first that $n_0 = 1$ if $h > 1$. In this case, it is easier to flip a negative spin surrounded by a positive spin than to flip a positive spin surrounded by two negative spins and the proof presented below simplifies. We assume that $h < 1$ so that $n_0 \geq 2$.

Recall the definition of the set $F_1(\sigma)$ introduced at the beginning of the proof. By the harmonic property of f_β,

$$f_\beta(\sigma) = \frac{1}{|F_1(\sigma)|} \sum_{\xi \in F_1(\sigma)} f_\beta(\xi) + \frac{e^{-2\beta[1-h]}}{|F_1(\sigma)|^2} \sum_{\eta \in G_1(\sigma)} \sum_{\xi \in F_1(\sigma)} [f_\beta(\eta) - f_\beta(\xi)] + o(\beta),$$

where $G_1(\sigma)$ is the set of configurations obtained from σ by flipping a negative spin surrounded by a positive spin and where $o(\beta)$ has an expression which vanishes faster than $e^{-2\beta[1-h]}$ as $\beta \uparrow \infty$.

We claim that

$$\lim_{\beta \to \infty} e^{2\beta[1-h]} \sum_{\xi \in F_1(\sigma)} \{f_\beta(\xi) - f_\beta(\sigma)\} = 0. \tag{8.3}$$

To prove this claim, denote by $F_k(\sigma)$, $1 \leq k \leq n_0$, the configurations obtained from σ by successively flipping k distinct positive spins surrounded by two negative spins: $F_{j+1}(\sigma) = \cup_{\xi \in F_j(\sigma)} F_1(\xi)$. Denote by $G_1(\eta)$ the predecessors of η, that is, the configurations obtained from η by flipping a negative spin surrounded by two positive spins. Hence, $G_1(\eta) \subset F_{j-1}(\sigma)$ if η belongs to $F_j(\sigma)$. By the harmonic property, for every $\eta \in F_j(\sigma)$, $1 \leq j < n_0$,

$$\sum_{\xi \in G_1(\eta)} \{f_\beta(\eta) - f_\beta(\xi)\} = e^{-\beta h} \sum_{\zeta \in F_1(\eta)} \{f_\beta(\zeta) - f_\beta(\eta)\} + O(e^{-\beta[2-h]}).$$

Replacing this identity in the sum appearing in (8.3), we reduce the proof of (8.3) to the proof that

$$e^{2\beta[1-h]} e^{-(n_0-1)\beta h} \sum_{\xi_1 \in F_1(\sigma)} \sum_{\xi_2 \in F_1(\xi_1)} \cdots \sum_{\xi_{n_0} \in F_1(\xi_{n_0-1})} \{f_\beta(\xi_{n_0}) - f_\beta(\xi_{n_0-1})\}$$

vanishes as $\beta \uparrow \infty$. This holds because f_β is bounded by 1 and $2/h < n_0 + 1$.

By the harmonic property of f_β at $\xi \in F_1(\sigma)$, $f(\xi) = f(\sigma)$ for any limit point f of the sequence f_β. Moreover, by (8.3) and by the displayed formula appearing just before (8.3),

$$\sum_{\eta \in G_1(\sigma)} [f(\eta) - f(\sigma)] = 0.$$

Recall the notation introduced in Section 3. Note that $G_1(\sigma) = \mathbb{W}(\sigma)$ and that $f(\eta) = [j + f(\sigma)]/(j + 1)$ if η belongs to $\mathbb{W}_j(\sigma, \sigma'), 1 \leq j \leq 3$, while $f(\eta) = f(\sigma)/(j + 1)$ if $\eta \in \mathbb{W}_j(\sigma) \setminus \mathbb{W}_j(\sigma, \sigma')$. This observation permits us to conclude the proof of the lemma.   □

Recall the definition of the sets $\Omega_{o,k}, \ 1 \leq k \leq n_0$, and $\mathbb{S}(\sigma)$ introduced in Section 3.

Corollary 8.3. *Fix a configuration σ in $\Omega_{o,k} \setminus \Omega_{o,k+1}, \ 1 \leq k \leq n_0, \sigma \neq +\mathbf{1}, -\mathbf{1}$. Let $\Omega_{k,\sigma} = \Omega_{o,k} \setminus \{\sigma\}$. Then, for all $\sigma' \in \Omega_{k,\sigma}$,*

$$\lim_{\beta \to \infty} \mathbf{P}_\sigma^\beta \big[T_{\sigma'} = T_{\Omega_{k,\sigma}} \big] = p(\sigma, \sigma').$$

Proof. Fix $1 \leq k \leq n_0$ and a configuration σ in $\Omega_{o,k} \setminus \Omega_{o,k+1}, \sigma \neq +\mathbf{1}, -\mathbf{1}$. For $k = 1$ and $k = n_0$, since $\mathbb{D}(\sigma) = \mathbb{S}(\sigma)$ and $\sum_{\sigma' \in \mathbb{D}(\sigma)} q(\sigma, \sigma') = 1$, by Lemma 8.2,

$$\lim_{\beta \to \infty} \mathbf{P}_\sigma^\beta \big[T_{\mathbb{S}(\sigma)} = T_{\Omega_\sigma} \big] = 1.$$

Since $\mathbb{S}(\sigma) \subset \Omega_{k,\sigma} \subset \Omega_\sigma$ we may replace $T_{\mathbb{S}(\sigma)}$ by $T_{\Omega_{k,\sigma}}$ in the previous equation. The corollary follows now from Lemma 8.2 and the fact that $p = q$ for $k = 1$ and $k = n_0$.

Consider now the case $2 \leq k < n_0$. Fix a configuration $\sigma' \in \mathbb{D}(\sigma) \cap \Omega_{o,k} \subset \mathbb{S}(\sigma)$. Since $T_{\Omega_\sigma} \leq T_{\Omega_{k,\sigma}}$, and since $p(\sigma, \sigma^*) = q(\sigma, \sigma^*)$ for $\sigma^* \in \mathbb{S}(\sigma)$, by Lemma 8.2,

$$\liminf_{\beta \to \infty} \mathbf{P}_\sigma^\beta \big[T_{\sigma'} = T_{\Omega_{k,\sigma}} \big] \geq \lim_{\beta \to \infty} \mathbf{P}_\sigma^\beta \big[T_{\sigma'} = T_{\Omega_\sigma} \big] = p(\sigma, \sigma').$$

Fix now a configuration $\sigma' \in \mathbb{S}(\sigma) \setminus \mathbb{D}(\sigma)$. This configuration is obtained from σ by flipping all spins of a positive $\ell(\sigma) \times \ell(\sigma)$ square of σ. Denote by $\sigma_j, \ 1 \leq j \leq 4$, the four configurations obtained from σ by flipping all spins from one of the sides of this square. Of course,

$$\mathbf{P}_\sigma^\beta \big[T_{\sigma'} = T_{\Omega_{k,\sigma}} \big] \geq \sum_{j=1}^4 \mathbf{P}_\sigma^\beta \big[T_{\sigma'} = T_{\Omega_{k,\sigma}}, T_{\sigma_j} = T_{\Omega_\sigma} \big].$$

Since $T_{\Omega_\sigma} \leq T_{\Omega_{k,\sigma}}$ and $\sigma_j \notin \Omega_{k,\sigma}$, by the strong Markov property, the right hand side is equal to

$$\sum_{j=1}^4 \mathbf{P}_\sigma^\beta \big[T_{\sigma_j} = T_{\Omega_\sigma} \big] \mathbf{P}_{\sigma_j}^\beta \big[T_{\sigma'} = T_{\Omega_{k,\sigma}} \big].$$

By Lemma 8.2, $\mathbf{P}_\sigma^\beta \big[T_{\sigma_j} = T_{\Omega_\sigma} \big]$ converges to $q(\sigma, \sigma_j)$ as $\beta \uparrow \infty$. We also claim that $\mathbf{P}_{\sigma_j}^\beta \big[T_{\sigma'} = T_{\Omega_{k,\sigma}} \big]$ converges to 1 as $\beta \uparrow \infty$ for $1 \leq j \leq 4$. Indeed, for a fixed j, $\ell(\sigma_j) = \ell(\sigma) - 1$ and the configuration σ_j has one and only one positive rectangle R with a side of length $\ell(\sigma) - 1$. It follows from Lemma 8.2 and from the definition of the sets $\mathbb{D}(\sigma^*)$ that the process first flips the spins of one side of the rectangle R, transforming it into a positive $[\ell(\sigma) - 1] \times [\ell(\sigma) - 1]$ square. Then, it flips the spins of one side of this square, transforming it into a positive $[\ell(\sigma) - 2] \times [\ell(\sigma) - 1]$ rectangle, and so on, until the process reaches a configuration where the initial rectangle R is transformed into a 2×2 square, without flipping in this process any other site which is not contained in the original $\ell(\sigma) \times \ell(\sigma)$ positive square of σ. In the last step, all spins of the 2×2 positive square are flipped and the process reaches the configuration σ' which belongs to $\Omega_{k,\sigma}$ and is the first one to belong to this set in the evolution just described. This proves the claim.

It follows from this argument that

$$\liminf_{\beta \to \infty} \mathbf{P}_\sigma^\beta \big[T_{\sigma'} = T_{\Omega_{k,\sigma}}\big] \geq \sum_{j=1}^4 q(\sigma, \sigma_j) = p(\sigma, \sigma').$$

Since this inequality holds for all $\sigma' \in \mathbb{S}(\sigma)$ and $\sum_{\sigma' \in \mathbb{S}(\sigma)} p(\sigma, \sigma') = 1$, the lemma is proved. $\square$

The proof of Lemma 8.2 describes the asymptotic behavior of $\mathbf{P}_\eta^\beta[T_\sigma < T_{\Omega_\sigma}]$ for some configurations η, but not for all. We may not, therefore, apply blindly Lemma 4.2 to deduce the limit of the capacity $\text{cap}_\beta(\{\sigma\}, \Omega_\sigma)$. The next result fills the gaps.

For $3 \leq \ell(\sigma) \leq n_0$, denote by $\mathbb{W}_1(\sigma)$ the configurations in $\mathbb{W}(\sigma)$ whose positive spin surrounded by three negative spins is in the corner of a positive rectangle of σ and denote by $\mathbb{W}_2(\sigma)$ the remaining configurations of $\mathbb{W}(\sigma)$. Note that configurations in $\mathbb{W}_j(\sigma)$ jump to Ω_σ with probability $(j+1)^{-1} + o(\beta)$ and that $|\mathbb{W}_1(\sigma)| = 4N_r(\sigma) + 8N_s(\sigma)$, $|\mathbb{W}_2(\sigma)| = 2[\ell(\sigma) - 2]N_r(\sigma) + 4[\ell(\sigma) - 2]N_s(\sigma)$.

Lemma 8.4. *Fix a configuration σ in Ω_o, $\sigma \neq +\mathbf{1}, -\mathbf{1}$. If $2 \leq \ell := \ell(\sigma) \leq n_0$,*

$$\lim_{\beta \to \infty} e^{\beta[\ell-1]h} \mu_\beta(\sigma)^{-1} \text{cap}_\beta(\{\sigma\}, \Omega_\sigma) = \theta(\sigma),$$

and if $\ell > n_0$,

$$\lim_{\beta \to \infty} e^{\beta(2-h)} \mu_\beta(\sigma)^{-1} \text{cap}_\beta(\{\sigma\}, \Omega_\sigma) = \theta(\sigma),$$

where $\theta(\sigma)$ has been defined in (3.4).

Proof. Fix a configuration σ satisfying the assumptions of the lemma and assume that $3 \leq \ell := \ell(\sigma) \leq n_0$. By Lemmas 4.1 and 8.1, we know that $\text{cap}_\beta(\{\sigma\}, \Omega_\sigma)$ is of order $\mu_\beta(\sigma)e^{-\beta(\ell-1)h}$.

We start with the proof of the upper bound for the capacity. Recall that we denote by $\mathbb{W}(\sigma)$ the set of saddle configurations of the valley $(\{\sigma\}, \{\sigma\} \cup \Delta, \sigma)$. Denote by B the set of all configurations η which do not belong to $\mathbb{W}(\sigma)$ and which can be reached from σ by self-avoiding paths $\gamma = (\sigma = \eta_0, \eta_1, \ldots, \eta_p = \eta)$ such that $\mu_\beta(\eta_k) \geq \mu_\beta(\sigma)e^{-\beta[\ell-1]h}$, $0 \leq k \leq p$. It follows from the proof of Lemma 8.1 that all these configurations are obtained from σ by successively flipping at most $\ell - 1$ positive spins which are surrounded by two negative spins. Note that all neighbors ξ of a configuration $\eta \in B$ which do not belong to B have measure $\mu_\beta(\xi) < \mu_\beta(\sigma)e^{-\beta[\ell-1]h}$.

Consider the function $f : \Omega \to [0, 1]$ defined as follows. Set $f(\sigma) = 1$, $f = 1$ on B, $f = j/(j+1)$ on $\mathbb{W}_j(\sigma)$ and $f = 0$ elsewhere. By the definition of capacity and by the definition of the function f, $\text{cap}_\beta(\{\sigma\}, \Omega_\sigma) \leq D_\beta(f) = \mu_\beta(\sigma)e^{-\beta[\ell-1]h}\{(2/3)|\mathbb{W}_2(\sigma)| + (1/2)|\mathbb{W}_1(\sigma)|\} + o(\beta)$, where $o(\beta) \prec \mu_\beta(\sigma)e^{-\beta[\ell-1]h}$. This proves the upper bound.

To prove the lower bound, consider a function f equal to 1 at σ and 0 on Ω_σ. Denote by A_0 the set of all configurations η which can be reached from σ by self-avoiding paths $\gamma = (\sigma = \eta_0, \eta_1, \ldots, \eta_p = \eta)$ such that $\mu_\beta(\eta_k) \geq \mu_\beta(\sigma)e^{-\beta[\ell-1]h}$, $0 \leq k \leq p$, and let $A = A_0 \cup \mathbb{D}(\sigma)$. By the definition of the Dirichlet form,

$$D_\beta(f) \geq \sum_{\{\eta, \xi\} \subset A} G_\beta(\eta, \xi)[f(\xi) - f(\eta)]^2.$$

Denote by $f_\beta : A \to [0, 1]$ the function which minimizes the right hand side with the boundary conditions imposed above. It is well known that $f_\beta(\eta) = \mathbf{P}_\eta^{A,\beta}[T_\sigma < T_{\mathbb{D}(\sigma)}]$ where $\mathbf{P}^{A,\beta}$ stands

for the probability on the path space induced by the reversible Markov process whose Dirichlet form is the one appearing on the right hand side of the previous formula. The asymptotic behavior of $f_\beta(\eta)$, as $\beta \uparrow \infty$, has been examined in the previous lemma for certain configurations. The arguments presented in the proof of the lower bound of Lemma 4.2 permit us to conclude.

The proofs for $\ell(\sigma) = 2 \le n_0$ and $\ell(\sigma) > n_0$ are simpler and left to the reader. $\qquad\square$

Recall the definition of the set $\Omega_{k,\sigma}$ introduced in Corollary 8.3 and fix $2 \le k \le n_0$. Since $\Omega_{k,\sigma} \subset \Omega_\sigma$, $\mathrm{cap}_\beta(\{\sigma\}, \Omega_{k,\sigma}) \le \mathrm{cap}_\beta(\{\sigma\}, \Omega_\sigma)$. The method of the proof of the lower bound for $\mathrm{cap}_\beta(\{\sigma\}, \Omega_\sigma)$, together with the asymptotic behavior of the hitting times stated in Corollary 8.3, provides the next result.

Corollary 8.5. *Fix* $2 \le k < n_0$ *and a configuration* σ *in* $\Omega_{k,\sigma} \setminus \Omega_{k+1,\sigma}$. *Then,*

$$\lim_{\beta\to\infty} \mathrm{e}^{\beta k h}\, \mu_\beta(\sigma)^{-1}\, \mathrm{cap}_\beta(\{\sigma\}, \Omega_{k,\sigma}) = \theta(\sigma).$$

Moreover, for σ *in* $\Omega_{n_0,\sigma} \setminus \Omega_{n_0+1,\sigma}$,

$$\lim_{\beta\to\infty} \mathrm{e}^{\beta(2-h)}\, \mu_\beta(\sigma)^{-1}\, \mathrm{cap}_\beta(\{\sigma\}, \Omega_{n_0,\sigma}) = \theta(\sigma).$$

9. Proofs of Theorems 3.1 and 3.2

The proof of Theorem 3.1 is based on the theory developed in the early sections of this article. A simple computation shows that $\mathbb{H}(\sigma^x) - \mathbb{H}(\sigma) = \sum_{y:|y-x|=1} \sigma(y)\,\sigma(x) + h\,\sigma(x)$, where $|\cdot|$ stands for the Euclidean norm. Since $0 < h < 2$, the jump rates $c(x,\sigma)$ may only assume the values 1, $\mathrm{e}^{-\beta[4+h]}$, $\mathrm{e}^{-\beta[2+h]}$, $\mathrm{e}^{-\beta h}$, $\mathrm{e}^{-\beta[4-h]}$ and $\mathrm{e}^{-\beta[2-h]}$. Assumptions (2.1), (2.2) are therefore satisfied.

Recall the terminology and the notation introduced in Section 5. According to the theory developed in the previous sections, the first step in the proof of the metastable behavior of a Markov process is the description of the evolution among the shallowest valleys which we now determine. Since a negative (resp. positive) spin surrounded by two (resp. three) positive (resp. negative) spins flips at rate 1, it is not difficult to show that the leaves $\mathcal{E}_1, \dots, \mathcal{E}_\nu$ defined in Section 5 are all the singletons formed by the elements of Ω_o and so $\nu = |\Omega_o|$ and $\Delta = \Omega \setminus \Omega_o$.

Denote by $\mathcal{E}_\sigma$ the singleton $\{\sigma\}$, $\sigma \in \Omega_o$. By Lemma 5.4, Proposition 5.7 and Lemma 8.4, $(\{\sigma\}, \{\sigma\} \cup \Delta, \sigma)$, $\sigma \ne -\mathbf{1}, +\mathbf{1}$, is a valley of depth $\mathrm{e}^{\beta[\ell(\sigma)-1]h}\theta(\sigma)^{-1}$ if $2 \le \ell(\sigma) \le n_0$ and of depth $\mathrm{e}^{\beta(2-h)}\theta(\sigma)^{-1}$ if $\ell(\sigma) > n_0$. Moreover, by Lemmas 8.1 and 4.1 and the same results invoked above, $(\{\pm\mathbf{1}\}, \{\pm\mathbf{1}\} \cup \Delta, \pm\mathbf{1})$ is a valley whose depth is of order $\mathrm{e}^{\beta(8\pm3h)}$. The exact depth of these latter valleys is not important at this stage.

To describe the evolution among the shallowest valleys, recall the notation introduced in Section 6. For a subset F of Ω, denote by $R_\beta^F(\sigma, \sigma')$, $\sigma, \sigma' \in F$, the jump rates of the trace σ_t^F of the process σ_t on F. Let $r_\beta^F(A, B)$, $A, B \subset F$, $A \cap B = \varnothing$, be the average rates of jumping of σ_t^F from A to B:

$$r_\beta^F(A, B) = \frac{1}{\mu_\beta(A)} \sum_{\sigma \in A} \mu_\beta(\sigma) \sum_{\sigma' \in B} R_\beta^F(\sigma, \sigma').$$

In view of the depths of the valleys $(\{\sigma\}, \{\sigma\} \cup \Delta, \sigma)$, $\sigma \in \Omega_o$, the set S_1 can be identified as the set $\Omega_{o,1} \setminus \Omega_{o,2}$. Recall that $\theta_\beta(1) = \mathrm{e}^{\beta h}$ if $n_0 \ge 2$ and $\theta_\beta(1) = \mathrm{e}^{\beta(2-h)}$ if $n_0 = 1$. By Lemma 6.1, Corollary 8.3, the explicit expression for the depth of the valleys obtained

above, and Lemma 10.2, the scaled average rates $e^{\beta h} r_\beta^{\Omega_o}(\sigma, \sigma')$, σ, $\sigma' \in \Omega_o$, converge to $r(\sigma, \sigma') = \theta(\sigma) p(\sigma, \sigma')$, where $p(\sigma, \sigma')$ and $\theta(\sigma)$ have been introduced in (3.1)–(3.4).

Recall that we denote by $\sigma_t^{\beta,1}$ the trace of the Markov process σ_t^β on $\Omega_{o,1}$. By Lemma 6.2 with $\theta_\beta(1) = e^{\beta h}$ and by the observations of the previous paragraph, the speeded-up process $\sigma_{t\theta_\beta(1)}^{\beta,1}$ converges to a Markov process on $\Omega_{o,1}$ with jump rates $r(\sigma, \sigma') = \theta(\sigma) p(\sigma, \sigma')$. By Proposition 6.3, on the time scale $\theta_\beta(1)$, the time spent in Δ is negligible. This proves Theorem 3.1 for $k = 1$.

The proof of Theorem 3.1 in the longer time scales is based on Theorem 7.1 and follows the strategy presented in Remark 7.2. Recall the notation introduced in Section 7 and assumption T. Since Theorem 3.1 has been proven for $k = 1$, assumption T holds at level 1 because all 1-metastates are singletons.

Theorem 3.1 for $2 \le k \le n_0$ follows from Theorem 7.1. As explained in Remark 7.2, we just need to characterize the metastates at each level, the depth of each valley and the asymptotic rates. This has been done for $2 \le k \le n_0$ in Corollaries 8.3 and 8.5, in view of Lemma 10.2. We present in detail the case $k = 2$ and leave the rest of the recursive argument to the reader.

Assume that $n_0 \ge 2$. It follows from the dynamics generated by the rates r introduced above that the leaves at level 2, $\mathcal{E}_1^{(2)}, \ldots, \mathcal{E}_{\nu(2)}^{(2)}$, are all the singletons formed by the elements of $\Omega_{o,2}$ so that $\nu(2) = |\Omega_{o,2}|$ and $\Delta_2 = \Omega \setminus \Omega_{o,2}$, $\Delta_2^o = \Omega_{o,1} \setminus \Omega_{o,2}$.

By Theorem 7.1 with $\mathfrak{p} = 1$ and Corollary 8.5, the triples $(\{\sigma\}, \{\sigma\} \cup \Delta_2^o, \sigma)$, $\sigma \in \Omega_{o,2}$, $\sigma \ne -\mathbf{1}, +\mathbf{1}$, are valleys for the trace process $\sigma_t^{\beta,1}$ of depth $e^{\beta[\ell(\sigma)-1]h} \theta(\sigma)^{-1}$ if $3 \le \ell(\sigma) \le n_0$ and of depth $e^{\beta(2-h)} \theta(\sigma)^{-1}$ if $\ell(\sigma) > n_0$. Moreover, by Lemmas 8.1 and 4.1, $(\{\pm\mathbf{1}\}\{\pm\mathbf{1}\} \cup \Delta_2^o, \pm\mathbf{1})$ is a valley for the trace process $\sigma_t^{\beta,1}$ whose depth is of order $e^{\beta(8\pm 3h)}$.

Note that the Ising model presents the particularity that the $\mathfrak{p}$-metastates are 1-metastates, and not a union of 1-metastates.

Recall the definition of the set S_2 introduced just after Lemma 7.10. The set S_2 can be identified as the set $\Omega_{o,2} \setminus \Omega_{o,3}$. Set $\theta_\beta(2) = e^{2\beta h}$ if $n_0 > 2$ and $\theta_\beta(2) = e^{\beta(2-h)}$ if $n_0 = 2$. By Theorem 7.1 with $\mathfrak{p} = 1$, Lemma 10.2, Corollaries 8.3 and 8.5, $\sigma_{t\theta_\beta(2)}^{\beta,2}$ converges to a Markov process on $\Omega_{o,2}$ with jump rates $r(\sigma, \sigma') = \theta(\sigma) p(\sigma, \sigma')$ introduced in (3.1)–(3.4). Furthermore, by Theorem 7.1 with $\mathfrak{p} = 1$, on the time scale $\theta_\beta(2)$, the time spent in Δ_2 is negligible. This proves Theorem 3.1 for $k = 2$. $\square$

We now turn to the proof of Theorem 3.2. It relies on the following lemma. Recall the definition of the sets $\mathbb{W}(-\mathbf{1})$, $\mathbb{W}_1(-\mathbf{1})$ and $\mathbb{W}_2(-\mathbf{1})$ and of the number $\theta(-\mathbf{1})$ introduced just before the statement of Theorem 3.2.

Lemma 9.1. *For $\beta > 0$, $G_\beta(\{-\mathbf{1}\}, \{+\mathbf{1}\}) = \mu_\beta(\sigma^*)$, for any $\sigma^* \in \mathbb{W}(-\mathbf{1})$. Moreover,*

$$\lim_{\beta \to \infty} e^{\beta c(h)} \mu_\beta(-\mathbf{1})^{-1} \mathrm{cap}_\beta(\{-\mathbf{1}\}, \{+\mathbf{1}\}) = \theta(-\mathbf{1}),$$

where $c(h) = 4(n_0 + 1) - h[(n_0 + 1)n_0 + 1]$ and $\theta(-\mathbf{1}) = (2/3)|\mathbb{W}_2(-\mathbf{1})| + (1/2)|\mathbb{W}_1(-\mathbf{1})|$.

Proof. The proof of the first assertion is left to the reader. The proof of the second one is similar to that of Lemma 4.2.

Denote by B the set of all configurations η which do not belong to $\mathbb{W}(-\mathbf{1})$ and which can be reached from $-\mathbf{1}$ by self-avoiding paths $\gamma = (-\mathbf{1} = \eta_0, \eta_1, \ldots, \eta_p = \eta)$ such that $\mu_\beta(\eta_k) \ge \mu_\beta(\sigma^*)$, $0 \le k \le p$, for some $\sigma^* \in \mathbb{W}(-\mathbf{1})$. All these configurations are obtained from $-\mathbf{1}$ by flipping at most $n_0(n_0+1)$ negative spins. Note that all neighbors ξ of a configuration $\eta \in B$ which do not belong to B have measure $\mu_\beta(\xi) < \mu_\beta(\sigma^*)$.

Consider the function $f : \Omega \to [0, 1]$ defined as follows. Set $f(-\mathbf{1}) = 1$, $f = 1$ on B, $f = 1/(j + 1)$ on $\mathbb{W}_j(\sigma)$ and $f = 0$ elsewhere. By the definition of capacity and by the definition of the function f, $\mathrm{cap}_\beta(\{-\mathbf{1}\}, \{+\mathbf{1}\}) \leq D_\beta(f) = \mu_\beta(\sigma^*)\{(2/3)|\mathbb{W}_2(-\mathbf{1})| + (1/2)|\mathbb{W}_1(-\mathbf{1})|\} + o(\beta)$, where $o(\beta) \prec \mu_\beta(\sigma^*)$. This proves the upper bound.

To prove the lower bound, recall that the function $f : \Omega \to \mathbb{R}$ which minimizes the Dirichlet form under the constraint that $f(-\mathbf{1}) = 1$, $f(+\mathbf{1}) = 0$ is the hitting time $g_\beta(\sigma) = \mathbf{P}^\beta_\sigma[T_{-\mathbf{1}} < T_{+\mathbf{1}}]$.

Denote by A the set of all neighbors ξ of $\mathbb{W}(-\mathbf{1})$ which are obtained from a configuration $\eta \in \mathbb{W}(-\mathbf{1})$ by either flipping the positive spin surrounded by three negative spins or by flipping a negative spin surrounded by two positive spins. By the definition of the Dirichlet form,

$$\mathrm{cap}_\beta(\{-\mathbf{1}\}, \{+\mathbf{1}\}) = D_\beta(g_\beta) \geq \sum_{\eta \in \mathbb{W}(-\mathbf{1}), \xi \in A} G_\beta(\eta, \xi)[g_\beta(\xi) - g_\beta(\eta)]^2.$$

It follows from Corollary 8.3 that $g_\beta(\xi)$ converges to 0 (resp. 1) as $\beta \uparrow \infty$ if ξ is a configuration obtained from a configuration in $\mathbb{W}(-\mathbf{1})$ by flipping a negative spin surrounded by two positive spins (resp. by flipping the positive spin surrounded by three negative spins). On the other hand, since g_β is harmonic and since a configuration $\eta \in \mathbb{W}_j(-\mathbf{1})$ jumps to configurations in A, where the asymptotic behavior of g_β is known, at rates of order 1, and jumps to other configurations at rate $o(\beta)$, $g_\beta(\eta)$ converges, as $\beta \uparrow \infty$, to $1/(j + 1)$ if $\eta \in W_j(-\mathbf{1})$. This proves the lower bound since $G_\beta(\eta, \xi) = \mu_\beta(\sigma^*)$ for $\eta \in \mathbb{W}(-\mathbf{1})$, $\xi \in A$. $\square$

We are now in a position to prove Theorem 3.2 which relies on Theorem 7.1 and the strategy presented in Remark 7.2. Up to this point we have proved assumption T at level n_0. In view of the asymptotic dynamics of the trace process σ_t^{β, n_0} described in Theorem 3.1, there are only two $(n_0 + 1)$-metastates, $\{-\mathbf{1}\}$ and $\{+\mathbf{1}\}$. By Theorem 7.1 and by Lemma 9.1, $(\{-\mathbf{1}\}, \{-\mathbf{1}\} \cup \Delta^o_{n_0+1}, -\mathbf{1})$ is a valley for the trace process σ_t^{β, n_0} of depth $e^{\beta c(h)}\theta(-\mathbf{1})^{-1}$. A computation similar to that presented in Lemma 9.1 shows that $(\{+\mathbf{1}\}, \{+\mathbf{1}\} \cup \Delta^o_{n_0+1}, +\mathbf{1})$ is a valley for the trace process σ_t^{β, n_0} whose depth is of magnitude larger than that of the valley $(\{-\mathbf{1}\}, \{-\mathbf{1}\} \cup \Delta^o_{n_0+1}, -\mathbf{1})$. Recall that $\theta_\beta(n_0 + 1) = e^{\beta c(h)}$ and note that we may identify the set S_{n_0+1} with the singleton $\{-\mathbf{1}\}$.

Since the state space of the trace process σ_t^{β, n_0+1} is a pair, by Theorem 7.1, by the explicit computation of the depth of the valley $(\{-\mathbf{1}\}, \{-\mathbf{1}\} \cup \Delta^o_{n_0+1}, -\mathbf{1})$ and by Lemma 10.2, the speeded-up trace process $\sigma_{t\theta_\beta(n_0+1)}^{\beta, n_0+1}$ converges to the Markov process on $\{-\mathbf{1}, +\mathbf{1}\}$ in which $+\mathbf{1}$ is an absorbing state and which jumps from $-\mathbf{1}$ to $+\mathbf{1}$ at rate $\theta(-\mathbf{1})$. The second assertion of Theorem 3.2 also follows from Theorem 7.1. $\square$

10. General results

We state in this section some general results on metastability of continuous time Markov chains used in the previous sections. We assume that the reader is familiar with the notation and terminology of [2].

Fix a sequence $(E_N : N \geq 1)$ of countable state spaces. The elements of E_N are denoted by the Greek letters η, ξ. For each $N \geq 1$ consider a matrix $R_N : E_N \times E_N \to \mathbb{R}$ such that $R_N(\eta, \xi) \geq 0$ for $\eta \neq \xi$, $-\infty < R_N(\eta, \eta) \leq 0$ and $\sum_{\xi \in E_N} R_N(\eta, \xi) = 0$ for all $\eta \in E_N$.

Let $\{\eta_t^N : t \geq 0\}$ be the *minimal* right-continuous Markov process associated with the jump rates $R_N(\eta, \xi)$ [22]. It is well known that $\{\eta_t^N : t \geq 0\}$ is a strong Markov process with respect

to the filtration $\{\mathcal{F}_t^N : t \geq 0\}$ given by $\mathcal{F}_t^N = \sigma(\eta_s^N : s \leq t)$. Let $\mathbf{P}_\eta$, $\eta \in E_N$, be the probability measure on $D(\mathbb{R}_+, E_N)$ induced by the Markov process $\{\eta_t^N : t \geq 0\}$ starting from η.

Consider two sequences $\mathscr{W} = (W_N \subseteq E_N : N \geq 1)$, $\mathscr{B} = (B_N \subseteq E_N : N \geq 1)$ of subsets of E_N, the second one containing the first and being properly contained in $E_N : W_N \subseteq B_N \subsetneq E_N$. Fix a point $\boldsymbol{\xi} = (\xi_N \in W_N : N \geq 1)$ in $\mathscr{W}$ and a sequence of positive numbers $\boldsymbol{\theta} = (\theta_N : N \geq 1)$.

Lemma 10.1. *Assume that the triple $(\mathscr{W}, \mathscr{B}, \boldsymbol{\xi})$ is a valley of depth $\boldsymbol{\theta}$ with attractor $\boldsymbol{\xi}$. Let $\mathscr{C} = (C_N \subset E_N : N \geq 1)$ be a sequence of sets such that:*
(1) $C_N \cap B_N = \varnothing$.
(2) For every $\delta > 0$,

$$\lim_{N \to \infty} \sup_{\eta \in C_N} \mathbf{P}_\eta \left[\frac{1}{\theta_N} T_{(\mathscr{B} \cup \mathscr{C})^c} > \delta \right] = 0. \tag{10.1}$$

Then, the triple $(\mathscr{W}, \mathscr{B} \cup \mathscr{C}, \boldsymbol{\xi})$ is a valley of depth $\boldsymbol{\theta}$ with attractor $\boldsymbol{\xi}$.

Proof. We need to check the three conditions of [2, Definition 2.1]. Since $(B_N \cup C_N)^c \subset B_N^c$, condition (V1) is clearly fulfilled.

To prove (V3), decompose the event $T_{(\mathscr{B} \cup \mathscr{C})^c}(\boldsymbol{\Delta} \cup \mathscr{C}) > \delta\theta_N$ according to whether $T_\mathscr{C} < T_{(\mathscr{B} \cup \mathscr{C})^c}$ or $T_\mathscr{C} > T_{(\mathscr{B} \cup \mathscr{C})^c}$. In the latter case, $T_{(\mathscr{B} \cup \mathscr{C})^c}(\boldsymbol{\Delta} \cup \mathscr{C}) = T_{\mathscr{B}^c}(\boldsymbol{\Delta})$, so for every point $\boldsymbol{\eta} = (\eta^N : N \geq 1)$ in $\mathscr{W}$,

$$\limsup_{N \to \infty} \mathbf{P}_{\eta^N} \left[\frac{1}{\theta_N} T_{(\mathscr{B} \cup \mathscr{C})^c}(\boldsymbol{\Delta} \cup \mathscr{C}) > \delta, \ T_\mathscr{C} > T_{(\mathscr{B} \cup \mathscr{C})^c} \right] \leq \lim_{N \to \infty} \mathbf{P}_{\eta^N} \left[\frac{1}{\theta_N} T_{\mathscr{B}^c}(\boldsymbol{\Delta}) > \delta \right]$$
$$= 0,$$

where the last identity follows from the fact that the triple $(\mathscr{W}, \mathscr{B}, \boldsymbol{\xi})$ is a valley and from condition (V3) in the definition of a valley. On the other hand, since on the set $T_\mathscr{C} < T_{(\mathscr{B} \cup \mathscr{C})^c}$,

$$T_{(\mathscr{B} \cup \mathscr{C})^c}(\boldsymbol{\Delta} \cup \mathscr{C}) = \int_0^{T_\mathscr{C}} \mathbf{1}\{\eta_s^N \in \Delta_N \cup C_N\} \, ds + \int_{T_\mathscr{C}}^{T_{(\mathscr{B} \cup \mathscr{C})^c}} \mathbf{1}\{\eta_s^N \in \Delta_N \cup C_N\} \, ds$$
$$= \int_0^{T_{\mathscr{B}^c}} \mathbf{1}\{\eta_s^N \in \Delta_N\} \, ds + \int_{T_\mathscr{C}}^{T_{(\mathscr{B} \cup \mathscr{C})^c}} \mathbf{1}\{\eta_s^N \in \Delta_N \cup C_N\} \, ds,$$

by the strong Markov property,

$$\mathbf{P}_{\eta^N} \left[\frac{1}{\theta_N} T_{(\mathscr{B} \cup \mathscr{C})^c}(\boldsymbol{\Delta} \cup \mathscr{C}) > \delta, \ T_\mathscr{C} < T_{(\mathscr{B} \cup \mathscr{C})^c} \right] \leq \mathbf{P}_{\eta^N} \left[\frac{1}{\theta_N} T_{\mathscr{B}^c}(\boldsymbol{\Delta}) > \delta/2 \right]$$
$$+ \sup_{\eta \in C_N} \mathbf{P}_\eta \left[\frac{1}{\theta_N} \int_0^{T_{(\mathscr{B} \cup \mathscr{C})^c}} \mathbf{1}\{\eta_s^N \in \Delta_N \cup C_N\} \, ds > \delta/2 \right].$$

The right hand side of this inequality vanishes as $N \uparrow \infty$ by hypothesis (10.1) and by the fact that the triple $(\mathscr{W}, \mathscr{B}, \boldsymbol{\xi})$ is a valley.

Putting together the two previous estimates, we obtain that for every $\delta > 0$ and every point $\boldsymbol{\eta} = (\eta^N : N \geq 1)$ in $\mathscr{W}$,

$$\lim_{N \to \infty} \mathbf{P}_{\eta^N} \left[\frac{1}{\theta_N} T_{(\mathscr{B} \cup \mathscr{C})^c}(\boldsymbol{\Delta} \cup \mathscr{C}) > \delta \right] = 0.$$

This shows that the triple $(\mathscr{W}, \mathscr{B} \cup \mathscr{C}, \boldsymbol{\xi})$ satisfies assumption (V3) for a valley with depth θ_N.

It remains to check that the assumption (V2) for a valley is fulfilled. On the one hand, since $T_{(\mathscr{B}\cup\mathscr{C})^c} \geq T_{\mathscr{B}^c}$ and since the triple $(\mathscr{W}, \mathscr{B}, \boldsymbol{\xi})$ is a valley of depth θ_N, for every $t > 0$ and every point $\boldsymbol{\eta} = (\eta^N : N \geq 1)$ in $\mathscr{W}$,

$$\liminf_{N\to\infty} \mathbf{P}_{\eta^N}\left[\frac{1}{\theta_N}T_{(\mathscr{B}\cup\mathscr{C})^c} > t\right] \geq \lim_{N\to\infty} \mathbf{P}_{\eta^N}\left[\frac{1}{\theta_N}T_{\mathscr{B}^c} > t\right] = e^{-t}. \tag{10.2}$$

On the other hand, decompose the set $\{T_{(\mathscr{B}\cup\mathscr{C})^c} > t\theta_N\}$ according to the partition $T_{\mathscr{C}} < T_{(\mathscr{B}\cup\mathscr{C})^c}$, $T_{\mathscr{C}} > T_{(\mathscr{B}\cup\mathscr{C})^c}$. In the latter set, $T_{(\mathscr{B}\cup\mathscr{C})^c} = T_{\mathscr{B}^c}$, while in the first one, $T_{(\mathscr{B}\cup\mathscr{C})^c} = T_{\mathscr{B}^c} + T_{(\mathscr{B}\cup\mathscr{C})^c} \circ T_{\mathscr{C}}$. Therefore, for every $t > 0$ and every point $\boldsymbol{\eta} = (\eta^N : N \geq 1)$ in $\mathscr{W}$,

$$\mathbf{P}_{\eta^N}\left[\frac{1}{\theta_N}T_{(\mathscr{B}\cup\mathscr{C})^c} > t\right] = \mathbf{P}_{\eta^N}\left[\frac{1}{\theta_N}T_{\mathscr{B}^c} > t, T_{\mathscr{C}} > T_{(\mathscr{B}\cup\mathscr{C})^c}\right]$$
$$+ \mathbf{P}_{\eta^N}\left[T_{\mathscr{B}^c} + T_{(\mathscr{B}\cup\mathscr{C})^c} \circ T_{\mathscr{C}} > t\theta_N, T_{\mathscr{C}} < T_{(\mathscr{B}\cup\mathscr{C})^c}\right].$$

By the strong Markov property, the second term on the right hand side is bounded above by

$$\sup_{\eta\in C_N} \mathbf{P}_{\eta}\left[T_{(\mathscr{B}\cup\mathscr{C})^c} > \delta\theta_N\right] + \mathbf{P}_{\eta^N}\left[T_{\mathscr{B}^c} > (t-\delta)\theta_N, T_{\mathscr{C}} < T_{(\mathscr{B}\cup\mathscr{C})^c}\right]$$

for every $\delta > 0$. Therefore, in view of the two previous displayed formulas, for every $\delta > 0$,

$$\mathbf{P}_{\eta^N}\left[\frac{1}{\theta_N}T_{(\mathscr{B}\cup\mathscr{C})^c} > t\right] \leq \mathbf{P}_{\eta^N}\left[\frac{1}{\theta_N}T_{\mathscr{B}^c} > t - \delta\right] + \sup_{\eta\in C_N} \mathbf{P}_{\eta}\left[T_{(\mathscr{B}\cup\mathscr{C})^c} > \delta\theta_N\right].$$

By (10.1), the second term on the right hand side vanishes as $N \uparrow \infty$ for every $\delta > 0$. Since the triple $(\mathscr{W}, \mathscr{B}, \boldsymbol{\xi})$ is a valley of depth θ_N, by condition (V2) for a valley, the first term converges to $e^{-(t-\delta)}$ as $N \uparrow \infty$. Hence, letting $\delta \downarrow 0$ after $N \uparrow \infty$, we obtain that for every $t > 0$ and every point $\boldsymbol{\eta} = (\eta^N : N \geq 1)$ in $\mathscr{W}$,

$$\limsup_{N\to\infty} \mathbf{P}_{\eta^N}\left[\frac{1}{\theta_N}T_{(\mathscr{B}\cup\mathscr{C})^c} > t\right] \leq e^{-t}.$$

This estimate together with (10.2) shows that the triple $(\mathscr{W}, \mathscr{B} \cup \mathscr{C}, \boldsymbol{\xi})$ satisfies condition (V2) for a valley with depth θ_N. $\square$

Of course, this result is only interesting if the process may jump from B_N to C_N.

10.1. The positive recurrent reversible case

We assume from now on that the Markov process $\{\eta_t^N : t \geq 0\}$ is positive recurrent and reversible with respect to its unique invariant probability measure denoted by μ_N.

Fix $N \geq 1$ and a proper subset F_N of E_N. Denote by $R^{F_N}(\eta, \xi)$ the jump rates of the trace of the process $\{\eta_t^N : t \geq 0\}$ on the set F_N. We refer the reader to [2, Section 6.1] for a precise definition. For each pair A, B of disjoint subsets of F_N, denote by $r_{F_N}(A, B)$ the average rate at which the trace process on F_N jumps from A to B:

$$r_{F_N}(A, B) = \frac{1}{\mu_N(A)} \sum_{\eta\in A} \mu_N(\eta) \sum_{\xi\in B} R^{F_N}(\eta, \xi).$$

We claim that

$$r_{F_N}(A, B) \leq \frac{\mathrm{cap}_N(A, B)}{\mu_N(A)}, \qquad (10.3)$$

where $\mathrm{cap}_N(A, B)$ stands for the capacity between A and B for the process $\{\eta_t^N : t \geq 0\}$. Indeed, denote by $R^{A \cup B}$ the jump rates of the trace of $\{\eta_t^N : t \geq 0\}$ on $A \cup B$. By [2, Corollary 6.2], $R^{A \cup B}(\eta, \xi) \geq R^{F_N}(\eta, \xi)$ for every $\eta, \xi \in A \cup B, \eta \neq \xi$. Hence, by definition of the average rates and by [2, Lemma 6.7],

$$r_{F_N}(A, B) = \frac{1}{\mu_N(A)} \sum_{\eta \in A} \mu_N(\eta) \sum_{\xi \in B} R^{F_N}(\eta, \xi) \leq r_{A \cup B}(A, B) = \frac{\mathrm{cap}_N(A, B)}{\mu_N(A)},$$

which proves (10.3).

Fix a finite number of disjoint subsets $\mathscr{E}_N^1, \ldots, \mathscr{E}_N^\kappa$, $\kappa \geq 2$, of $E_N : \mathscr{E}_N^x \cap \mathscr{E}_N^y = \varnothing$, $x \neq y$. Let $\mathscr{E}_N = \cup_{x \in S} \mathscr{E}_N^x$ and let $\breve{\mathscr{E}}_N^x := \mathscr{E}_N \setminus \mathscr{E}_N^x$. Denote by $r_N(\mathscr{E}_N^x, \mathscr{E}_N^y)$ the average rates $r_{\mathscr{E}_N}(\mathscr{E}_N^x, \mathscr{E}_N^y)$ and denote by $P_N(\eta, x, y)$, $1 \leq x \neq y \leq \kappa$, $\eta \in \mathscr{E}_N^x$, the probabilities of hitting among metastates:

$$P_N(\eta, x, y) := \mathbf{P}_\eta[T_{\mathscr{E}^y} = T_{\breve{\mathscr{E}}^x}].$$

The next result shows that if the average rates appropriately rescaled converge, their limit can be expressed in terms of the depth of the metastates and their hitting probabilities.

Lemma 10.2. *Suppose that for each $1 \leq x \leq \kappa$ there exists a point $\boldsymbol{\xi}_x = (\xi_x^N : N \geq 1)$ in $\mathscr{E}^x$ such that the triple $(\mathscr{E}^x, \mathscr{E}^x \cup \boldsymbol{\Delta}, \boldsymbol{\xi}_x)$ is a valley of depth $\Lambda_N(x)^{-1}$ and such that*

$$\lim_{N \to \infty} \sup_{\eta \in \mathscr{E}_N^x} \frac{\mathrm{cap}_N(\mathscr{E}^x, \breve{\mathscr{E}}^x)}{\mathrm{cap}_N(\xi_x^N, \eta)} = 0.$$

Suppose, furthermore, that there exists a sequence $(\theta_N : N \geq 1)$ for which the mean rates, the depth and the jump probabilities converge: for any $1 \leq x \neq y \leq \kappa$ and any sequence $(\eta^N : N \geq 1)$ in $\mathscr{E}^x$,

$$\lim_{N \to \infty} \theta_N \Lambda_N(x) = \Lambda(x), \qquad \lim_{N \to \infty} P_N(\eta^N, x, y) = P(x, y)$$
$$\lim_{N \to \infty} \theta_N r_N(\mathscr{E}^x, \mathscr{E}^y) = r(x, y).$$

Then, $r(x, y) = \Lambda(x) P(x, y)$.

Proof. Note that we assumed that the limit $P(x, y)$ does not depend on the sequence $(\eta^N : N \geq 1)$.

It follows from [2, Theorem 2.7] that for any $1 \leq x \leq \kappa$ and any sequence $(\eta^N : N \geq 1)$ in $\mathscr{E}^x$, under the measure $\mathbf{P}_{\eta^N}$ the speeded-up process $X_t^N = \Psi(\eta_{t\theta_N}^N)$ converges to a Markov process on $\{1, \ldots, \kappa\}$ with jump rates $r(y, z)$, starting from x. In particular, if we denote by τ_1^N the time of the first jump of X_t^N, τ_1^N converges to an exponential time of rate $\lambda(x) = \sum_{y \neq x} r(x, y)$ and $X_{\tau_1^N}^N$ converges to a random variable with distribution $p(y) = r(x, y)/\lambda(y)$.

On the other hand, since the triple $(\mathscr{E}^x, \mathscr{E}^x \cup \boldsymbol{\Delta}, \boldsymbol{\xi}_x)$ is a valley of depth $\Lambda_N(x)^{-1}$, $\tau_1^N \Lambda_N(x)$ converge to a mean 1 exponential time and so $\Lambda(x) = \lambda(x)$. Moreover, $\mathbf{P}_{\eta^N}[X_{\tau_1^N}^N = y] = \mathbf{P}_{\eta^N}[T_{\mathscr{E}^y} = T_{\breve{\mathscr{E}}^x}]$ converges to $P(x, y)$, so $P(x, y) = p(y)$, which proves the lemma. $\square$

Acknowledgement

References

[1]   J. Beltrán, C. Landim, Metastability of reversible condensed zero range processes on a finite set, 2009. arXiv:0910.4089.

[2]   J. Beltrán, C. Landim, Tunneling and metastability of continuous time Markov chains, J. Stat. Phys. 140 (2010) 1065–1114.

[3]   G. Ben Arous, R. Cerf, Metastability of the three-dimensional Ising model on a torus at very low temperature, Electron. J. Probab. 1 (1996) Research Paper 10.

[4]   A. Bovier, Metastability: a potential theoretic approach, in: International Congress of Mathematicians, vol. III, Eur. Math. Soc., Zürich, 2006, pp. 499–518.

[5]   A. Bovier, M. Eckhoff, V. Gayrard, M. Klein, Metastability in stochastic dynamics of disordered mean field models, Probab. Theory Related Fields 119 (2001) 99–161.

[6]   A. Bovier, M. Eckhoff, V. Gayrard, M. Klein, Metastability and low lying spectra in reversible Markov chains, Comm. Math. Phys. 228 (2002) 219–255.

[7]   A. Bovier, F. Manzo, Metastability in Glauber dynamics in the low-temperature limit: beyond exponential asymptotics, J. Stat. Phys. 107 (2002) 757–779.

[8]   M. Cassandro, A. Galves, E. Olivieri, M.E. Vares, Metastable behavior of stochastic dynamics: a pathwise approach, J. Stat. Phys. 35 (1984) 603–634.

[9]   F. Den Hollander, Metastability under stochastic dynamics, Stochastic Process. Appl. 114 (2004) 1–26.

[10] F. Den Hollander, Three lectures on metastability under stochastic dynamics, in: R. Kotecký (Ed.), Methods of Contemporary Mathematical Statistical Physics, in: Lecture Notes in Math., vol. 1970, Springer, Berlin, 2009.

[11] F. den Hollander, E. Olivieri, E. Scoppola, Metastability and nucleation for conservative dynamics, J. Math. Phys. 41 (2000) 1424–1498.

[12] F. den Hollander, E. Olivieri, E. Scoppola, Nucleation in fluids: some rigorous results, Physica A 279 (2000) 110–122.

[13] F. Den Hollander, F.R. Nardi, E. Olivieri, E. Scoppola, Droplet growth for three-dimensional Kawasaki dynamics, Probab. Theory Related Fields 125 (2003) 153–194.

[14] A. Gaudillière, F. Den Hollander, F.R. Nardi, E. Olivieri, E. Scoppola, Ideal gas approximation for a two-dimensional rarefied gas under Kawasaki dynamics, Stochastic Process. Appl. 119 (2009) 737–774.

[15] M. Jara, C. Landim, A. Teixeira, Quenched scaling limits of trap models, Ann. Probab. 39 (2011) 176–223.

[16] M. Jara, C. Landim, A. Teixeira, Quenched scaling limits of trap models in random graphs. (in preparation).

[17] R. Kotecky, E. Olivieri, Droplet dynamics for asymmetric Ising model, J. Stat. Phys. 70 (1993) 1121–1148.

[18] R. Kotecky, E. Olivieri, Shapes of growing droplets–a model of escape from a metastable phase, J. Stat. Phys. 75 (1994) 409–506.

[19] F. Manzo, F. Nardi, E. Olivieri, E. Scoppola, On the essential features of metastability: tunnelling time and critical configurations, J. Stat. Phys. 115 (2004) 591–642.

[20] E.J. Neves, R.H. Schonmann, Critical droplets and metastability for a Glauber dynamics at very low temperatures, Comm. Math. Phys. 137 (1991) 209–230.

[21] E.J. Neves, R.H. Schonmann, Behavior of droplets for a class of Glauber dynamics at very low temperature, Probab. Theory Related Fields 91 (1992) 331–354.

[22] J.R. Norris, Markov Chains, Cambridge University Press, Cambridge, 1997.

[23] E. Olivieri, M.E. Vares, Large Deviations and Metastability, in: Encyclopedia of Mathematics and its Applications, vol. 100, Cambridge University Press, Cambridge, 2005.

[24] E. Scoppola, Renormalization group for Markov chains and application to metastability, J. Stat. Phys. 73 (1993) 83–121.