Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2019, Vol. 55, No. 4, 1850–1889
https://doi.org/10.1214/18-AIHP936

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Metastability of one-dimensional, non-reversible diffusions with periodic boundary conditions

C. Landim[a] and I. Seo[b]

[a]*IMPA, Estrada Dona Castorina 110, CEP 22460 Rio de Janeiro, Brasil and CNRS UMR 6085, Université de Rouen, France.*
E-mail: *landim@impa.br*
[b]*Department of Mathematical Sciences and Research Institute of Mathematics, Seoul National University, Seoul, Korea.*
E-mail: *insuk.seo@snu.ac.kr*

Received 18 October 2017; revised 24 August 2018; accepted 20 September 2018

Abstract. We consider small perturbations of a dynamical system on the one-dimensional torus. We derive sharp estimates for the pre-factor of the stationary state, we examine the asymptotic behavior of the solutions of the Hamilton–Jacobi equation for the pre-factor, we compute the capacities between disjoint sets, and we prove the metastable behavior of the process among the deepest wells following the martingale approach. We also present a bound for the probability that a Markov process hits a set before some fixed time in terms of the capacity of an enlarged process.

Résumé. Nous considérons de petites perturbations d'un système dynamique sur le tore unidimensionnel. Nous obtenons des estimations précises pour le pré-facteur de l'état stationnaire, nous examinons le comportement asymptotique des solutions de l'équation de Hamilton–Jacobi pour le pré-facteur, nous calculons les capacités entre des ensembles disjoints et nous prouvons le comportement métastable du processus parmi les puits les plus profonds en suivant l'approche martingale. Nous présentons également une borne pour la probabilité qu'un processus de Markov atteigne un ensemble avant un certain instant en termes de capacité d'un processus élargi.

MSC: 60J60; 60F99

Keywords: Non-reversible diffusions; Potential theory; Metastability; Dirichlet principle; Thomson principle; Eyring–Kramers formula

1. Introduction

Variational formulae for the capacity between two sets have been derived recently in the context of continuous time Markov chains and diffusions [9,13,18]. These formulae were used to prove the metastable behavior of asymmetric condensing zero-range processes [11], random walks in a potential field [15], mean field Potts model [14], and to derive the Eyring–Kramers formula for the transition time of non-reversible diffusions [13].

To estimate the capacity through the variational formulae alluded to above, one needs to know explicitly the stationary state. This property is shared by all the dynamics mentioned in the previous paragraph, where the invariant measures are the equilibrium states of the reversible version of the dynamics. Usually, however, the stationary states of non-reversible Markovian dynamics are not known explicitly.

It is possible, nonetheless, to derive through dynamical large deviations methods a formula for the quasi-potential of non-reversible dynamics and estimates for the stationary state with exponentially small errors [8]. A natural development of this approach consists in using potential theory to get sharper bounds of the stationary state, that is, to provide precise estimates for the first-order term in the expansion of the quasi-potential, the so-called pre-factor.

For one-dimensional diffusion processes with periodic boundary conditions,

$$dX_\varepsilon(t) = b\big(X_\varepsilon(t)\big)\,dt + \sqrt{2\varepsilon}\,dW_t, \tag{1.1}$$

where $b : \mathbb{T} \to \mathbb{R}$ is a smooth drift, $\varepsilon > 0$ a small parameter and W_t the Brownian motion on the one-dimensional torus $\mathbb{T} = [0, 1)$, one may derive sharp estimates for the pre-factor due to an explicit formula for the stationary state obtained by Faggionato and Gabrielli [7]. This estimate and the precise bounds for the capacity between two wells constitute the first main result of the article.

The pre-factor of the stationary state solves a Hamilton–Jacobi equation. We take advantage of the explicit formulae to examine the asymptotic behavior of the solution of the Hamilton–Jacobi equation in the hope that these results might give some insight on the behavior of the pre-factor in higher dimensions.

The second main result of this article provides an extension to diffusion processes of the martingale approach proposed in [1–3] to derive the metastable behavior of Markov chains. The main difficulty in applying this method to diffusions lies in the fact that the martingale approach requires an analysis of the trace of the process on the wells. While for Markov chains the trace process is still a Markov chain [with long range jumps], for diffusions the trace becomes a singular diffusion with jumps along the boundary of the wells, a dynamics very different from the original one and difficult to analyze.

We present in this article an entirely new approach inspired by results in partial differential equations obtained by Evans, Tabrizian and Seo, Tabrizian [6,17]. Here is the idea. Denote by $\mathcal{E}_i$, $1 \le i \le n$, the wells, and let G be a function defined on the entire space and which is constant [with possibly different values] in each well. Denote by L_ε the generator and by F_ε the solution of the Poisson equation $L_\varepsilon F_\varepsilon = G$. Assume that for all such functions G the solution F_ε is uniformly bounded and is asymptotically constant in each well. We prove in Section 7 that the convergence in law of the projection of the trace process on the wells follows from the previous property of the solutions of the Poisson equation.

This new way of deriving the metastable behavior of a Markov chain is applied here to small perturbations of the dynamical system (1.1). It also provides the first example where the reduced chain, which describes the asymptotic dynamics among the wells, is a irreducible, non-reversible Markov chain.

The third main result of the article consists in a bound on the probability that the hitting time of a set is less than or equal to a constant in terms of capacities. In view of the variational formulae for the capacity, this result provides a general method to obtain upper bounds for the probability of an event which appears in many different contexts.

We conclude this introduction with some historical remarks and a description of the article. The convergence of the order parameter, in the sense of finite-dimensional distributions, of small perturbations by reversible Gaussian noises of dynamical systems has been proved by Sugiura in [19]. Imkeller and Pavlyukevich [10] obtained a similar result in one-dimension when the Brownian motion is replaced by a Lévy process. More recently, Bouchet and Reygner [4] derived a formula for the transition time between two wells for non-reversible diffusions. A rigorous proof of this result is still an open problem.

The paper is organized as follows. In Section 2, we introduce the model and the main assumptions on the drift b. In Section 3, we present the main results of the article in the case of two wells. In Section 4, we introduce the notion of valleys and landscapes used throughout the article. In Sections 5 and 6, we derive sharp asymptotic estimates for the pre-factor and for the capacity between two wells. In Section 7, we prove the metastable behavior of the process by showing that the projection of the trace process on the wells converges in an appropriate time scale to a finite-state Markov chain. In Section 8, we prove a bound on the probability that a certain set is attained before a fixed time in terms of the capacities of an enlarged process. Finally, in Section 9, we prove that the solutions of certain Poisson equations are asymptotically constant on the wells.

2. The model

We introduce in this section the model, the main assumptions and known results.

2.1. *The diffusion process*

Let $\mathbb{T} = [0, 1)$ be the one-dimensional torus of length 1. Consider a continuous vector field $b : \mathbb{T} \to \mathbb{R}$. Throughout this article, we assume that b fulfills the following conditions:

(H1) The closed set $\{\theta \in \mathbb{T} : b(\theta) = 0\}$ has a finite number of connected components, denoted by $I_j = [l_j, r_j]$, $1 \le j \le p$, for $0 < l_1 \le r_1 < l_2 \le \cdots < l_p \le r_p < 1$. Some of these intervals may be degenerate, as we do not exclude the possibility that $l_j = r_j$.
(H2) b is of class C^2 in the set $\mathbb{T} \setminus I$, where $I = \bigcup_{1\le j\le p} I_j$.
(H3) If (c, d) is a connected component of $\mathbb{T} \setminus I$, then $b'(c+) \neq 0$ and $b'(d-) \neq 0$. When the interval $[l_j, r_j]$ is degenerate, $l_j = r_j$, the left and right derivatives of b at r_j may be different: it may happen that $b'(r_j-) \neq b'(r_j+)$. However, both right and left derivatives do not vanish.

The generator of the diffusion (1.1), denoted by L_ε, is given by

$$(L_\varepsilon f)(\theta) = b(\theta) f'(\theta) + \varepsilon f''(\theta).$$

If the average drift vanishes,

$$\int_{\mathbb{T}} b(\theta)\, d\theta = 0,$$

there exists a potential $U : \mathbb{T} \to \mathbb{R}$ such that $b(\theta) = -U'(\theta)$. In this case the stationary measure is given by $Z_\varepsilon^{-1} \exp\{-U(\theta)/\varepsilon\}\, d\theta$ for a suitable normalization factor Z_ε^{-1} and the process is reversible with respect to this measure.

Assume, from now on, that

$$B := \int_{\mathbb{T}} b(\theta)\, d\theta > 0, \tag{2.1}$$

so that the process $X_\varepsilon(t)$ is non-reversible. In [7], the stationary measure of this process has been explicitly computed. Regard b as an 1-periodic function on $\mathbb{R}$. Let $S : \mathbb{R} \to \mathbb{R}$ be the function given by

$$S(x) = -\int_0^x b(z)\, dz, \tag{2.2}$$

and let $\pi_\varepsilon, m_\varepsilon : \mathbb{R} \to \mathbb{R}_+$ be given by

$$\pi_\varepsilon(x) = \int_x^{x+1} e^{[S(y)-S(x)]/\varepsilon}\, dy, \qquad m_\varepsilon(x) = \frac{1}{c(\varepsilon)} \pi_\varepsilon(x), \tag{2.3}$$

where $c(\varepsilon)$ is the normalizing constant which turns m_ε the density of a probability measure on $\mathbb{T}$. Indeed, $\pi_\varepsilon, m_\varepsilon$ are 1-periodic and can be considered as defined on $\mathbb{T}$. By [7], the measure $\mu_\varepsilon(d\theta) = m_\varepsilon(\theta)\, d\theta$ on $\mathbb{T}$ is the stationary state of the diffusion (1.1).

2.2. *The quasi-potential*

Let $z : \mathbb{R} \to \mathbb{R}$ be the function which indicates the position of the farthest maxima of S to the right: $z(x) = z_x$ is the largest point in $[x, \infty)$ at which a maximum of the set $\{S(y) : y \ge x\}$ is attained. More precisely,

(a) $S(z_x) = \max\{S(y) : y \ge x\}$,
(b) If $y \ge x$ and $S(y) = S(z_x)$, then $y \le z_x$.

Note that, for all $x \in \mathbb{R}$, $z(x)$ not only exists but also satisfies $z(x) \in [x, x+1)$ because $S(y+1) = S(y) - B < S(y)$. Moreover, z_x is a local maximum of S if $z_x \neq x$. In this case, $b(z_x) = -S'(z_x) = 0$.

Let $\widehat{V} : \mathbb{R} \to \mathbb{R}$ be given by

$$\widehat{V}(x) = S(x) - S(z_x). \tag{2.4}$$

Since $S(y + k) = S(y) - kB$ for $y \in \mathbb{R}$, $k \in \mathbb{Z}$, $z(x + k) = z(x) + k$. In particular, $\widehat{V}(x + 1) = \widehat{V}(x)$, so that $\widehat{V}$ is a 1-periodic function and can be considered as defined on the torus $\mathbb{T}$. By [7, Proposition 2.1], $\widehat{V}$ is the quasi-potential associated to the diffusion (1.1), and by [7, Theorem 2.4] it is a viscosity solution of the Hamilton–Jacobi equation associated to the Hamiltonian $H(x, p) = p[p - b(x)]$.

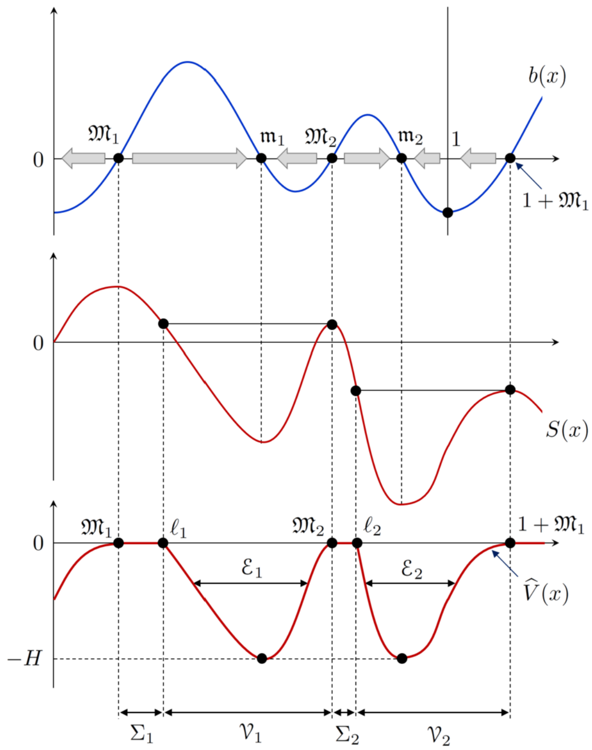

Fig. 1. The graphs of b, S and $\widehat{V}$. In the first graph, the gray arrows represent the direction of the drift in the dynamical system $dX(t) = b(X(t))\,dt$. Thus, $\mathfrak{m}_1$ and $\mathfrak{m}_2$ are stable equilibria, and $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are unstable equilibria. The existence of two stable equilibria separated by unstable one implies a metastable behavior of the perturbed dynamics (1.1).

3. Main result: Two stable points

We present in this section the main results of the article in the case, illustrated in Figure 1, where the drift b is smooth and the dynamical system $dX(t) = b(X(t))\,dt$ exhibits two stable equilibria and two unstable ones.

In addition to the conditions (H1)–(H3), we shall assume in this section that

(H0) The drift b is smooth and the set $\{\theta \in \mathbb{T} : b(\theta) = 0\}$ consists of four points.

Condition (H0) is not needed in the proofs of the results presented in this article. It is assumed in this section because it simplifies significantly the notation and the statement of the results, helping the reader to access the content of the article. Further assumptions will be formulated along the section.

3.1. *Notation*

By assumptions (H0) and (H3), $S(\cdot)$ has two local maxima $\mathfrak{M}_1$, $\mathfrak{M}_2$ and two local minima $\mathfrak{m}_1$, $\mathfrak{m}_2$. Without loss of generality, we assume that $0 < \mathfrak{M}_1 < \mathfrak{m}_1 < \mathfrak{M}_2 < \mathfrak{m}_2 < 1$, and that that $S(\mathfrak{M}_1) > S(\mathfrak{M}_2) > S(\mathfrak{M}_3)$, where we adopted the convention that $\mathfrak{M}_3 = 1 + \mathfrak{M}_1$. We refer to Figure 1 for the graphs of $b(\cdot)$, $S(\cdot)$ and $\widehat{V}(\cdot)$.

For each $i = 1, 2$, let $\ell_i = \inf\{x > \mathfrak{M}_i : S(x) = S(\mathfrak{M}_{i+1})\}$ and set

$$\Sigma_i = (\mathfrak{M}_i, \ell_i), \qquad \mathcal{V}_i = (\ell_i, \mathfrak{M}_{i+1}), \quad i = 1, 2.$$

Note that $\mathcal{V}_2$ can be regarded as the subset of $\mathbb{T}$ given by $(\ell_2, 1] \cup (0, \mathfrak{M}_1)$. These notations will be comprehensively extended to a general drift b in Section 4. Note that for $x \in \Sigma_1 \cup \Sigma_2$, one has that $z(x) = x$ and hence $\widehat{V}(x) = 0$ (cf. Figure 1). Thus, the sets Σ_i, $i = 1, 2$, represent the *saddle intervals* between two *valleys* $\mathcal{V}_1$ and $\mathcal{V}_2$. The notion of valley is extended to the one of *landscape* in Section 4 to handle more general situations. The depth of the valleys $\mathcal{V}_1$ and $\mathcal{V}_2$ are $-\widehat{V}(\mathfrak{m}_1)$ and $-\widehat{V}(\mathfrak{m}_2)$, respectively. Assume that

$$\widehat{V}(\mathfrak{m}_1) = \widehat{V}(\mathfrak{m}_2) = -H, \tag{3.1}$$

so that the depth of the two valleys coincide. This assumption is not necessary for the results below, but without it the results become trivial and can easily be deduced from the argument.

For $i = 1, 2$, let $\mathcal{E}_i = [e_i^-, e_i^+] \subseteq \mathcal{V}_i$ be such that $\widehat{V}(\mathfrak{m}_i) < \widehat{V}(e_i^-) = \widehat{V}(e_i^+) < 0$ so that $\mathfrak{m}_i \in \mathcal{E}_i$. The set $\mathcal{E}_i$ represents the metastable well around the stable point $\mathfrak{m}_i$.

3.2. *Sharp asymptotics for the pre-factor*

The first main result of the article provides a sharp estimate of the stationary state. Write m_ε as

$$m_\varepsilon(\theta) = F_\varepsilon(\theta) e^{-V(\theta)/\varepsilon},$$

where $V(\theta) = \widehat{V}(\theta) + H$. The function $F_\varepsilon(\cdot)$ is called the pre-factor. Its behavior as $\varepsilon \to 0$ plays a fundamental role in the estimation of the capacity between two wells, which is one of the crucial steps in the proof of the metastable behavior of a Markov chain. Such a result is still open in the non reversible context except in the trivial case where the pre-factor is constant. The first main result of this article provides an expansion in ε of the pre-factor. For $i = 1, 2$, let

$$\sigma(\mathfrak{m}_i) = \sqrt{\frac{2\pi}{-b'(\mathfrak{m}_i)}}, \qquad \omega(\mathfrak{M}_i) = \sqrt{\frac{2\pi}{b'(\mathfrak{M}_i)}}, \qquad Z = \sum_{j=1}^{2} \sigma(\mathfrak{m}_j)\omega(\mathfrak{M}_{j+1}).$$

Theorem 3.1. *Under the assumptions* (H0)–(H3),

1. (*Pre-factor on valleys*) *for all* $x \in \mathcal{V}_i$, $i = 1, 2$,

$$F_\varepsilon(x) = \big[1 + o(1)\big] \frac{\omega(\mathfrak{M}_{i+1})}{Z\sqrt{\varepsilon}}$$

where $o(1) \to 0$ *as* $\varepsilon \to 0$ *uniformly on* $\mathcal{V}_1 \cup \mathcal{V}_2$.

2. (*Pre-factor on saddle intervals*) *for all* $x \in \Sigma_1 \cup \Sigma_2$,

$$F_\varepsilon(x) = \big[1 + o(1)\big] \frac{1}{Zb(x)}.$$

where $o(1) \to 0$ *as* $\varepsilon \to 0$ *for all* $x \in \Sigma_1 \cup \Sigma_2$.

The general case, without assumption (H0), is presented in Propositions 5.2 and 5.3. Note the difference in the scaling factor in parts (1) and (2). This difference is explained along with a connection to the Hamilton–Jacobi equation for F_ε in Section 5.5. The scaling difference of the pre-factor indicates that its asymptotic analysis in higher dimension may be a difficult problem.

3.3. *Metastable behavior*

We turn to the metastable behavior of the diffusion $X_\varepsilon(t)$ between the valleys $\mathcal{E}_1$ and $\mathcal{E}_2$. Let $\widehat{X}_\varepsilon(t) := X_\varepsilon(e^{H/\varepsilon}t)$ be the speeded-up process. As in the approach developed in [1,2], we define the metastable behavior of the diffusion as the convergence of the projection of the trace process.

To define the trace process of $\widehat{X}_\varepsilon(t)$ on $\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_2$, let

$$T_{\mathcal{E}}(t) = \int_0^t \chi_{\mathcal{E}}\big(\widehat{X}_\varepsilon(s)\big)\, ds, \qquad S_{\mathcal{E}}(t) = \sup\big\{s \geq 0 : T_{\mathcal{E}}(s) \leq t\big\}.$$

In this formula and below, χ_A, $A \subset \mathbb{R}$, represents the indicator function of the set A:

$$\chi_A(x) = 1 \quad \text{for } x \in A, \qquad \chi_A(x) = 0 \quad \text{otherwise}.$$

The process $Y_\varepsilon(t) = \widehat{X}_\varepsilon(S_{\mathcal{E}}(t))$ is called the trace of the process $\widehat{X}_\varepsilon$ on $\mathcal{E}$. Informally, one obtains a trajectory of $Y_\varepsilon(t)$ from $\widehat{X}_\varepsilon(t)$ by deleting the excursion of $\widehat{X}_\varepsilon$ outside $\mathcal{E}$. In Section 7.4, we show that $Y_\varepsilon(\cdot)$ is a Markov process on $\mathcal{E}$ with respect to a suitable filtration. Let $\Psi : \mathcal{E} \to \{1, 2\}$ be the projection defined by $\Psi(x) = \chi_{\mathcal{E}_1}(x) + 2\chi_{\mathcal{E}_2}(x)$. Clearly, $\boldsymbol{x}_\varepsilon(t) = \Psi(Y_\varepsilon(t))$ takes values in the set $\{1, 2\}$, and represents the valley visited by the process $Y_\varepsilon(t)$. Following [1,2], we shall say that the process $X_\varepsilon(t)$ is metastable in the time-scale $e^{H/\varepsilon}$ if $\boldsymbol{x}_\varepsilon(\cdot)$ converges to a Markov chain on $\{1, 2\}$, and if the process $\widehat{X}_\varepsilon(t)$ remains outside $\mathcal{E}$ for a negligible amount of time.

The method developed in [1–3] provides a robust way to establish these results. Moreover, it has been shown in [12] that under some mild extra assumptions the metastability as stated above entails the convergence of the finite-dimensional distributions of the projections of $\widehat{X}_\varepsilon(t)$.

This approach to metastability was successfully enforced in the context of Markov chains. Its extension to diffusions, such as the one considered in the current paper, faced a major difficulty due to the singular behavior of the trace process $Y_\varepsilon(t)$ at the boundary of $\mathcal{E}$. In this paper, we propose a new way of establishing the convergence of the projection of the trace process based on results from the theory of partial differential equations (cf. [6,17]). This is the content of Sections 7 and 9.

Let

$$R(1,2) = \frac{1}{\sigma(\mathfrak{m}_1)\omega(\mathfrak{M}_2)}, \qquad R(2,1) = \frac{1}{\sigma(\mathfrak{m}_2)\omega(\mathfrak{M}_1)},$$

and consider the Markov chain $\boldsymbol{X}(t)$ on $\{1, 2\}$ with generator given by

$$(\boldsymbol{L} f)(i) = R(i,j)\big[F(j) - F(i)\big], \quad \{i,j\} = \{1,2\}.$$

Let $\boldsymbol{Q}_j$, $j = 1, 2$, be the law of the Markov chain $\boldsymbol{X}(t)$ starting from j.

Theorem 3.2. *Fix $j \in \{1, 2\}$, $\theta_0 \in \mathcal{E}_j$ and suppose that $X_\varepsilon(0) = \theta_0$ for all $\varepsilon > 0$. Then, the law of the process $Y_\varepsilon(\cdot)$ converges to $\boldsymbol{Q}_j$ as $\varepsilon \to 0$.*

The general version of this result is presented in Theorem 7.3. Although, under (H0), the process $\boldsymbol{X}(t)$ is reversible with respect to its invariant distribution, this is no longer true in the general setting. Actually, Theorem 7.3 provides the first example of a dynamics whose asymptotic evolution is described by a non-reversible and irreducible Markov chain.

The proof of Theorem 3.2 is divided in two parts. We have first to establish the tightness of the process $\boldsymbol{x}_\varepsilon(t)$ (cf. Section 7.5). The core step in the proof of this result is an estimation of the escape time from a metastable well. For this purpose we establish a general inequality, presented in Proposition 8.1, which bounds the hitting time of a set in terms of a capacity which can be easily estimated through the variational formulae for the capacity. We believe that this inequality, the third main result of the article, can be useful in numerous different contexts.

The second part of the proof consists in the characterization of the limit point. This part is based on the analysis of the solution of a certain Poisson equation (cf. Proposition 9.1). This sort of analysis has been carried out in [6,17] for reversible diffusions based on ideas from PDEs.

4. Valleys and landscapes

We introduce in this section the notion of valleys and landscapes which play an important role in the description of the quasi-potential $\widehat{V}$.

Let $\mathcal{A}_k = [\mathfrak{M}_k^-, \mathfrak{M}_k^+]$, $1 \le k \le q$, $q \ge 0$, (resp. $\mathcal{U}_j = [\mathfrak{m}_j^-, \mathfrak{m}_j^+]$, $1 \le j \le q'$, $q' \ge 0$) be the sub-intervals of $[0, 1)$ where S assumes a local maximum (resp. minimum). For the local maxima, this means that b vanishes on each interval $[\mathfrak{M}_k^-, \mathfrak{M}_k^+]$ and that $b'(\mathfrak{M}_k^+) > 0$, $b'(\mathfrak{M}_k^-) > 0$. Similar relations hold for the intervals where S attains a local minimum. Note that q, q' might be equal to 0. The set $\mathcal{A}_1$ is represented in Figure 2.

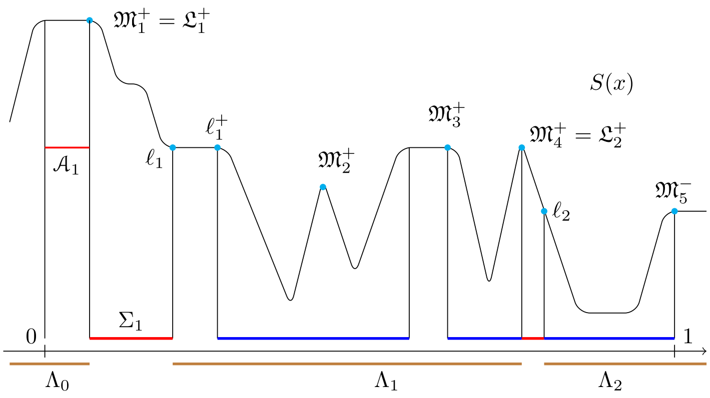

Fig. 2. This figure represents the graph on the interval $[-\delta, 1+\delta]$ of a function $S : \mathbb{R} \to \mathbb{R}$ associated by (2.2) to a vector field b defined on the torus $\mathbb{T}$. In this example, the set $\{\theta \in \mathbb{T} : b(\theta) = 0\}$ has $p = 10$ connected components. There are $q = 4$ local maxima, and $\boldsymbol{m} = 2$ landscapes indicated in brown. The landscape Λ_1 has two valleys represented in blue, while the landscape Λ_2 has only 1 valley. The two saddle intervals Σ_1, Σ_2 are displayed in red.

Since each local maxima is succeeded by a local minima, q' must be equal to q. The intervals $\mathcal{A}_k$, $\mathcal{U}_k$ might be reduced to points, and they are supposed to be ordered in the sense that $0 \le \mathfrak{M}_1^- \le \mathfrak{M}_1^+ < \mathfrak{m}_1^- \le \mathfrak{m}_1^+ < \mathfrak{M}_2^- < \cdots < \mathfrak{M}_q^+ < \mathfrak{m}_q^- \le \mathfrak{m}_q^+ < 1$. For $m \in \mathbb{Z}$, $1 \le j \le q$, let $\mathfrak{M}_{j+mq}^{\pm} = m + \mathfrak{M}_j^{\pm}$, $\mathfrak{m}_{j+mq}^{\pm} = m + \mathfrak{m}_j^{\pm}$.

Recall the definition of the function $z : \mathbb{R} \to \mathbb{R}$ introduced in Section 2.2. Observe that

$$z(x) = x \quad \text{or} \quad z(x) \in \left\{\mathfrak{M}_k^+ : \mathfrak{M}_k^+ > x\right\}. \tag{4.1}$$

Indeed, if $z(x) \ne x$, $z(x)$ must be greater than x and $z(x)$ must be the right endpoint of an interval where S attains a local maximum.

If $q = 0$, the diffusion $X_\varepsilon(t)$ has a nonnegative drift. In this case, S is a non-increasing function and $S(x) = S(z_x)$ for all $x \in \mathbb{R}$ so that the quasi-potential $\widehat{V}$, introduced in (2.4), vanishes: $\widehat{V} \equiv 0$. Conversely, if $\widehat{V} \equiv 0$, $S(z_x) = S(x)$ for all x, which implies that S has no local minima. Hence,

$$\widehat{V} \equiv 0 \quad \text{if and only if} \quad b \ge 0 \quad \text{if and only if} \quad q = 0.$$

Assume from now on that $q \ge 1$. Consider a local maximum $\mathfrak{M}_k^+$ such that

$$z\big(\mathfrak{M}_k^+\big) = \mathfrak{M}_k^+. \tag{4.2}$$

There is at least one maximum which comply with this condition: if $\mathfrak{M}_1^+$ does not fulfill it, then $z(\mathfrak{M}_1^+)$ does. In Figure 2, $\mathfrak{M}_1^+$, $\mathfrak{M}_4^+$ are the maxima which satisfy this condition.

Denote by $\boldsymbol{m}$ the number of local maxima which satisfy condition (4.2), and represent them by $\mathfrak{L}_n^{\pm} = \mathfrak{M}_{j(n)}^{\pm}$, $1 \le n \le \boldsymbol{m}$, for some sequence $1 \le j(1) < \cdots < j(\boldsymbol{m}) \le q$. As $\boldsymbol{m} \ge 1$ and $\boldsymbol{m} \le q$, we have that $0 \le \mathfrak{L}_1^+ < \cdots < \mathfrak{L}_{\boldsymbol{m}}^+ < 1$. Assume, without loss of generality, that $\mathfrak{L}_1^+ = \mathfrak{M}_1^+$, and extend the definition of the maxima $\mathfrak{L}_k^{\pm}$ by setting $\mathfrak{L}_{n+r\boldsymbol{m}}^{\pm} = r + \mathfrak{L}_n^{\pm}$, $r \in \mathbb{Z}$, $1 \le n \le \boldsymbol{m}$.

Assertion 4.A. *Fix a point $\mathfrak{L}_n^+ = \mathfrak{M}_{j(n)}^+$. We claim that $\mathfrak{L}_{n+1}^+ = z(\mathfrak{M}_{j(n)+1}^+)$.*

Proof. Figure 2 illustrates this assertion, as $\mathfrak{L}_2^+ = z(\mathfrak{M}_2^+)$. To prove the claim, we have to show that $z(\mathfrak{M}_{j(n)+1}^+)$ fulfills (4.2) and that no maxima $\mathfrak{M}_k^+$ in the interval $(\mathfrak{M}_{j(n)}^+, z(\mathfrak{M}_{j(n)+1}^+))$ satisfies (4.2). It is clear that $z(\mathfrak{M}_{j(n)+1}^+)$ fulfills (4.2). We turn to the second property. By (4.1), $z(\mathfrak{M}_{j(n)+1}^+) = \mathfrak{M}_\ell^+$ for some $\ell \ge j(n)+1$. By definition of z, $S(\mathfrak{M}_k^+) \le S(z(\mathfrak{M}_{j(n)+1}^+))$ or all $j(n)+1 \le k < \ell$. Hence $\mathfrak{M}_k^+$ does not satisfy (4.2) for k in this range, which proves the claim. $\square$

Fix a point $\mathfrak{L}_n^+ = \mathfrak{M}_{j(n)}^+$. Consider the next maximum, $\mathfrak{M}_{j(n)+1}^+$. As $z(\mathfrak{M}_{j(n)}^+) = \mathfrak{M}_{j(n)}^+$, $S(\mathfrak{M}_{j(n)+1}^+) < S(\mathfrak{M}_{j(n)}^+)$ and $S(\mathfrak{L}_{n+1}^+) = S(z(\mathfrak{M}_{j(n)+1}^+)) < S(\mathfrak{M}_{j(n)}^+)$. Let ℓ_n be the first point x larger than $\mathfrak{L}_n^+$ such that $S(x) = S(\mathfrak{L}_{n+1}^+)$:

$$\ell_n = \inf\bigl\{x \ge \mathfrak{L}_n^+ : S(x) = S\bigl(\mathfrak{L}_{n+1}^+\bigr)\bigr\}.$$

We refer to Figure 2 for a representation of ℓ_1 and ℓ_2. It is clear that $S(\ell_n) = S(\mathfrak{L}_{n+1}^+)$.

Assertion 4.B. *Fix a point $\mathfrak{L}_n^+ = \mathfrak{M}_{j(n)}^+$. Then,*

$$\widehat{V}(x) = \begin{cases} 0 & \mathfrak{L}_n^- \le x \le \ell_n, \\ S(x) - S(\mathfrak{L}_{n+1}^+) & \ell_n \le x \le \mathfrak{L}_{n+1}^+. \end{cases}$$

In particular, $\widehat{V}(\ell_n) = \widehat{V}(\mathfrak{L}_{n+1}^+) = 0$, and on the set $[\ell_n, \mathfrak{L}_{n+1}^+]$, the functions $\widehat{V}$ and S differ by an additive constant.

Proof. We leave to the reader to check that

$$S\bigl(z(x)\bigr) = \begin{cases} S(x) & \mathfrak{L}_n^- \le x \le \ell_n, \\ S(\mathfrak{L}_{n+1}^+) & \ell_n \le x \le \mathfrak{L}_{n+1}^+. \end{cases}$$

The assertion follows from this identity and from the definition of $\widehat{V}$. □

Note that it is not true that $z(x) = x$ for $x \in [\mathfrak{L}_n^+, \ell_n]$ because there might be subintervals of $[\mathfrak{L}_n^+, \ell_n]$ where S is constant. We refer to Figure 2. In contrast, $z(x) = \mathfrak{L}_{n+1}^+$ for $x \in [\ell_n, \mathfrak{L}_{n+1}^+]$.

The previous result characterizes the function $\widehat{V}$ in the interval $[\mathfrak{L}_n^-, \mathfrak{L}_{n+1}^+]$ and, therefore, in $\mathbb{T}$. Let $\ell_{r\boldsymbol{m}+n} = r + \ell_n$, $r \in \mathbb{Z}$, $1 \le n \le \boldsymbol{m}$, and let

$$\Sigma_n = \bigl(\mathfrak{L}_n^+, \ell_n\bigr), \qquad \Lambda_n = \bigl[\ell_n, \mathfrak{L}_{n+1}^+\bigr], \qquad \Sigma = \bigcup_{n=1}^{\boldsymbol{m}} \Sigma_n, \qquad \Lambda = \bigcup_{n=1}^{\boldsymbol{m}} \Lambda_n. \tag{4.3}$$

The sets Λ_n are named landscapes and the sets Σ_n saddle intervals. Of course, $\{\Sigma, \Lambda\}$ forms a partition of $\mathbb{T}$.

Remark 4.1. By Assertion 4.B and by definition of Λ_n, Σ_n, $\widehat{V}$ vanishes on Σ and $\widehat{V}$ and S differ by an additive constant on each landscape Λ_n. This additive constant may be different at each set Λ_n.

Note that $\mathfrak{L}_n^- < \ell_n$. Hence, even if the connected components of the set $\{\theta \in \mathbb{T} : b(\theta) = 0\}$ are points (that is, if $r_i = l_i$ for all $1 \le i \le p$), the intervals Σ_n at which the quasi-potential $\widehat{V}$ vanishes are non-degenerate. (See Figure 1).

By Assertion 4.B, $S(\ell_n) = S(\mathfrak{L}_{n+1}^+)$. If $\ell_n = l_k$ for some $1 \le k \le p$, let $\ell_n^+ = r_k$, otherwise let $\ell_n^+ = \ell_n$:

$$\ell_n^+ = \sup\bigl\{x \ge \ell_n : S(y) = S(\ell_n) \text{ for all } \ell_n \le y \le x\bigr\}.$$

Each landscape $\Lambda_n = [\ell_n, \mathfrak{L}_{n+1}^+]$ may contain in $(\ell_n^+, \mathfrak{L}_{n+1}^-)$ local maxima $\mathfrak{M}_k^+$ of S such that $S(\mathfrak{M}_k^+) = S(\ell_n)$ (and thus $\widehat{V}(\mathfrak{M}_k^+) = 0$). Let $\ell_n^+ < \mathfrak{M}_{a(n,1)}^+ < \cdots < \mathfrak{M}_{a(n,r_n-1)}^+ < \mathfrak{L}_{n+1}^-$ be an enumeration of these local maxima. Set $a(n, r_n) = j(n+1)$ so that $\mathfrak{M}_{a(n,r_n)}^+ = \mathfrak{M}_{j(n+1)}^+ = \mathfrak{L}_{n+1}^+$. The sets

$$\begin{aligned} &\mathcal{V}_{n,1} = \bigl(\ell_n^+, \mathfrak{M}_{a(n,1)}^-\bigr), \qquad \mathcal{V}_{n,r_n} = \bigl(\mathfrak{M}_{a(n,r_n-1)}^+, \mathfrak{L}_{n+1}^-\bigr), \\ &\mathcal{V}_{n,j+1} = \bigl(\mathfrak{M}_{a(n,j)}^+, \mathfrak{M}_{a(n,j+1)}^-\bigr), \quad 1 \le j \le r_n - 2 \end{aligned} \tag{4.4}$$

are called the valleys of the landscape Λ_n. To simplify some equations below let $\mathfrak{M}_{a(n,0)}^+ = \ell_n^+$, $\mathfrak{M}_{a(n,0)}^- = \ell_n^- = \ell_n$. Note that there is an abuse of notation since $\mathfrak{M}_{a(n,0)}^\pm = \ell_n^\pm$ may not be a maxima.

In Figure 2, the landscape $\Lambda_1 = [\ell_1, \mathfrak{M}_4^+]$ has 2 valleys, $(\ell_1^+, \mathfrak{M}_3^-)$ and $(\mathfrak{M}_3^+, \mathfrak{M}_4^-)$, while the landscape $\Lambda_2 = [\ell_2, \mathfrak{M}_5^+]$ has only one valley. Each landscape has at least one valley. If there are no local maxima $\mathfrak{M}_k^+$ of S in $(\ell_n^+, \mathfrak{L}_{n+1}^-)$ such that $S(\mathfrak{M}_k^+) = S(\ell_n)$, then $r_n = 1$ and the set $(\ell_n^+, \mathfrak{L}_{n+1}^-)$ forms a valley.

Remark 4.2. The quasi-potential $\widehat{V}$ may have plateaux $\{\theta \in \mathbb{T} : \widehat{V}(\theta) = 0\}$ which are not saddle intervals, but which belong to a landscape. This happens if one of the local maxima $\mathfrak{M}_{a(n,j)}^+$, $0 \le j \le r_n$, introduced above is such that $\mathfrak{M}_{a(n,j)}^+ \neq \mathfrak{M}_{a(n,j)}^-$ [with the convention adopted concerning $a(n,0)$ and $a(n,r_n)$]. This possibility is illustrated in Figure 2 by the intervals $[\mathfrak{M}_3^-, \mathfrak{M}_3^+]$, $[\ell_1, \ell_1^+]$. However, if the connected components of the set $\{\theta \in \mathbb{T} : b(\theta) = 0\}$ are points, all plateaux of $\widehat{V}$ are saddle intervals because in the landscapes the quasi-potential differ from S by an additive constant.

Remark 4.3. In a landscape $\Lambda_n = [\ell_n, \mathfrak{L}_{n+1}^+]$, the process $X_\varepsilon(t)$ evolves among the valleys $\mathcal{V}_{n,j}$ as a reversible process until it leaves Λ_n. Since $S(\mathfrak{L}_n^-) > S(\ell_n) = S(\mathfrak{L}_{n+1}^+)$, with a probability exponentially close to 1, $X_\varepsilon(t)$ leaves the landscape Λ_n through the saddle interval Σ_{n+1}. In Σ_{n+1}, as the drift is nonnegative, $X_\varepsilon(t)$ slides to the next landscape Λ_{n+1}. Once in Λ_{n+1}, with a probability exponentially close to 1, the process $X_\varepsilon(t)$ does not return to Σ_{n+1}. In particular, the saddle interval Σ_{n+1} is only visited during the excursion from Λ_n to Λ_{n+1}. This explains why the quasi-potential $\widehat{V}$ vanishes on the saddle intervals.

Remark 4.4. It is not possible to recover S from $\widehat{V}$. Given a maximal interval $[\theta_1, \theta_2]$ at which $\widehat{V}$ is constant equal to 0, it not possible to determine whether this interval is a saddle interval or whether it belongs to a landscape. However, if the connected components of $\{\theta \in \mathbb{T} : b(\theta) = 0\}$ are points, it is possible to recover S from $\widehat{V}$ and the pre-factor introduced in the next subsection.

5. The stationary state

One important question in the theory of non-reversible Markovian dynamics is to access the stationary state. Bounds for the quasi-potential with small exponential errors can be deduced from the theory of large deviations [8]. We present in Propositions 5.2, 5.3 and 5.7 below sharp asymptotics for the first-order term of the expansion in ε of the quasi-potential, the so-called pre-factor, defined in (5.2) below.

Precise estimates of the pre-factor play a central role in the derivation of the metastable behavior of a random process based on the potential theory, as one needs to evaluate the measure of a valley and the capacity between valleys (cf. [1,2,5]). An asymptotic analysis of the pre-factor for non-reversible dynamics similar to the one presented in this section has never been carried out before.

One available tool to obtain estimates for the pre-factor is the Hamilton–Jacobi equation (cf. (5.12) below). Write this equation as $H_\varepsilon(F_\varepsilon) = 0$. One is tempted to argue that F_ε should converge, as $\varepsilon \to 0$, to the solution of $H_0(F_0) = 0$. We show in Section 5.5 the limits of this analysis, proving that F_ε converges to a function which is discontinuous at the saddle points.

The main results of this section are based on the explicit expression (2.3) for the stationary state obtained by Faggionato and Gabrielli in [7]. Some of the claims below appear in [7]. They are stated here in sake of completeness as they will be used in the next sections.

5.1. *Definition*

Recall the definition of the quasi-potential $\widehat{V}$ and let $V : \mathbb{T} \to \mathbb{R}_+$ be the non-negative function given by

$$V(\theta) = \widehat{V}(\theta) + H. \tag{5.1}$$

Since the quasi-potential is defined up to constants, V can be regarded as another version of the quasi-potential. Write the density $m_\varepsilon(\theta)$ of the stationary distribution as

$$m_\varepsilon(\theta) = F_\varepsilon(\theta) e^{-V(\theta)/\varepsilon}. \tag{5.2}$$

The function F_ε is called the pre-factor, and corresponds to the first order correction of the quasi-potential.

5.2. *Sharp asymptotics*

We introduce three functions $G_k : \mathbb{T} \to \mathbb{R}$, $0 \le k \le 2$, which appear in the pre-factor. These functions are defined separately on each interval Λ_n, Σ_n, $1 \le n \le \boldsymbol{m}$.

We first consider the landscape. Fix $1 \le n \le \boldsymbol{m}$ and consider the set $\Lambda_n = [\ell_n, \mathfrak{L}^+_{n+1}]$. Denote by $G_0 : \Lambda_n \to \mathbb{R}_+$ the function given by

$$G_0(x) = \int_x^{\mathfrak{L}^+_{n+1}} \mathbf{1}\big\{S(y) = S(\ell_n)\big\}\, dy.$$

In this formula, $\mathbf{1}\{A\}$ takes the value 1 if A holds and 0 otherwise. The value of G_0 at x provides the Lebesgue measure of the set $[x, \mathfrak{L}^+_{n+1}] \cap \{y \in \mathbb{R} : S(y) = S(\ell_n)\}$. Note that G_0 is non-increasing, that it is constant on each valley of the landscape Λ_n, and that it vanishes if the connected components of $\{\theta \in \mathbb{T} : b(\theta) = 0\}$ are points.

We turn to the definition of G_1. Denote by $\mathfrak{n}_i^{\pm}$, $1 \le i \le q$, a local maximum $\mathfrak{M}_i^{\pm}$ or a local minimum $\mathfrak{m}_i^{\pm}$ of S, and let

$$\omega_+\big(\mathfrak{n}_i^+\big) = \sqrt{\frac{\pi}{2|b'(\mathfrak{n}_i^+ +)|}}, \qquad \omega_-\big(\mathfrak{n}_i^-\big) = \sqrt{\frac{\pi}{2|b'(\mathfrak{n}_i^- -)|}}. \tag{5.3}$$

Recall the definition of the valleys $\mathcal{V}_{n,j}$, $1 \le j \le r_n$, introduced in (4.4), and that $\mathfrak{M}^+_{a(n,0)} = \ell_n^+$. Denote by $G_1 : \Lambda_n \to \mathbb{R}_+$ the function given by

$$\begin{aligned} G_1(x) &= \omega_+\big(\mathfrak{M}^+_{a(n,0)}\big)\mathbf{1}\big\{b\big(\mathfrak{M}^+_{a(n,0)}\big) = 0, x \le \mathfrak{M}^+_{a(n,0)}\big\} \\ &\quad + \sum_{j=1}^{r_n}\big[\omega_-\big(\mathfrak{M}^-_{a(n,j)}\big)\mathbf{1}\big\{x < \mathfrak{M}^-_{a(n,j)}\big\} + \omega_+\big(\mathfrak{M}^+_{a(n,j)}\big)\mathbf{1}\big\{x \le \mathfrak{M}^+_{a(n,j)}\big\}\big]. \end{aligned} \tag{5.4}$$

Remark 5.1. The function G_1 is non-increasing. It may be discontinuous at $\mathfrak{M}^+_{a(n,0)}$, it is discontinuous at the points $\mathfrak{M}^-_{a(n,j)}$, $\mathfrak{M}^+_{a(n,j)}$, $1 \le j \le r_n$, and it is constant on the valleys $\mathcal{V}_{n,k} = (\mathfrak{M}^+_{a(n,k)}, \mathfrak{M}^-_{a(n,k+1)})$, $0 \le k < r_n$. Actually, we defined the valleys as open intervals instead of closed ones for the last property to hold.

We turn to the definition of the pre-factor on the saddle intervals. Fix $1 \le n \le \boldsymbol{m}$ and consider the set $\Sigma_n = (\mathfrak{L}^+_n, \ell_n)$. This set may contain connected components of the set $\{\theta \in \mathbb{T} : b(\theta) = 0\}$. Denote by $s_n \ge 0$ the number of such components and by $[c^-_{n,1}, c^+_{n,1}], \dots, [c^-_{n,s_n}, c^+_{n,s_n}]$ the components. Note that some of these intervals might be points: $c^-_{n,i}$ may be equal to $c^+_{n,i}$. Assume that these intervals are ordered in the sense that $c^-_{n,i} < c^-_{n,i+1}$.

Let

$$\mathscr{F}_n = \bigcup_{j=1}^{s_n} [c^-_{n,j}, c^+_{n,j}], \qquad \mathscr{G}_n = \Sigma_n \setminus \mathscr{F}_n, \qquad \mathscr{F} = \bigcup_{n=1}^{\boldsymbol{m}} \mathscr{F}_n, \qquad \mathscr{G} = \bigcup_{n=1}^{\boldsymbol{m}} \mathscr{G}_n. \tag{5.5}$$

Define $G_0 \colon \Sigma_n \to \mathbb{R}_+$ as

$$G_0(x) = z(x) - x.$$

As S is non-increasing on Σ_n, G_0 vanishes on $\mathscr{G}_n$ and $G_0(x) = c^+_{n,j} - x$ on the interval $[c^-_{n,j}, c^+_{n,j}]$:

$$G_0(x) = \sum_{j=1}^{s_n} \big(c^+_{n,j} - x\big)\chi_{[c^-_{n,j}, c^+_{n,j}]}(x).$$

Define $G_1 \colon \Sigma_n \to \mathbb{R}_+$ as

$$G_1(x) = \sum_{j=1}^{s_n} \omega_+\big(c^+_{n,j}\big)\chi_{[c^-_{n,j}, c^+_{n,j}]}(x).$$

As G_0, the function G_1 vanishes on on $\mathscr{G}_n$. Finally, define $G_2 \colon \Sigma_n \to \mathbb{R}_+$ as

$$G_2(x) = \frac{1}{b(x)} \chi_{\mathscr{G}_n}(x).$$

We are now in a position to present a sharp asymptotics for $\pi_\varepsilon(\cdot)$.

Proposition 5.2. *Assume that $q \geq 1$. Then,*

(1) (*Sharp estimate on the landscapes*)

$$\lim_{\varepsilon\to 0} \sup_{x\in\Lambda} \frac{1}{\sqrt{\varepsilon}} e^{\widehat{V}(x)/\varepsilon} \bigl|\pi_\varepsilon(x) - e^{-\widehat{V}(x)/\varepsilon}\bigl\{G_0(x) + \sqrt{\varepsilon} G_1(x)\bigr\}\bigr| = 0.$$

(2) (*Sharp estimate on the saddle intervals*) *On the set $\mathscr{F}$,*

$$\pi_\varepsilon(x) = \bigl\{G_0(x) + \sqrt{\varepsilon} G_1(x) + o(\sqrt{\varepsilon})\bigr\} e^{-\widehat{V}(x)/\varepsilon},$$

and on the set $\mathscr{G}$,

$$\pi_\varepsilon(x) = \bigl[1 + o(1)\bigr] \varepsilon G_2(x) e^{-\widehat{V}(x)/\varepsilon}.$$

In these formulas and below, $o(\sqrt{\varepsilon})$, resp. $o(1)$ represent quantities [*which may depend on x*] *with the property that* $\lim_{\varepsilon\to 0} o(\sqrt{\varepsilon})/\sqrt{\varepsilon} = 0$, *resp.* $\lim_{\varepsilon\to 0} o(1) = 0$.

We turn to the normalizing constant $c(\varepsilon)$. For $1 \leq k \leq q$, let $\sigma(\mathfrak{m}_k^+)$ be the weights given by

$$\sigma\bigl(\mathfrak{m}_k^+\bigr) = \omega_+\bigl(\mathfrak{m}_k^+\bigr) + \omega_-\bigl(\mathfrak{m}_k^-\bigr), \tag{5.6}$$

where the weights $\omega_\pm$ have been introduced in (5.3). Denote by H the depth of the deepest well,

$$H = -\min_{\theta\in\mathbb{T}} \widehat{V}(\theta) = \max_{0\leq x<1} \bigl\{S(z_x) - S(x)\bigr\}.$$

Clearly, $H > 0$ because $H \geq S(z(\mathfrak{m}_1^+)) - S(\mathfrak{m}_1^+) = S(z(\mathfrak{M}_2^+)) - S(\mathfrak{m}_1^+) \geq S(\mathfrak{M}_2^+) - S(\mathfrak{m}_1^+) > 0$. Let $\mathbb{I}$ be the set given by

$$\mathbb{I} = \bigl\{j \in \{1, \dots, q\} : \widehat{V}\bigl(\mathfrak{m}_j^+\bigr) = -H\bigr\}, \tag{5.7}$$

and let Z_ε be the normalizing constant given by

$$Z_\varepsilon = \sum_{j\in\mathbb{I}} \bigl\{G_0\bigl(\mathfrak{m}_j^+\bigr) + \sqrt{\varepsilon} G_1\bigl(\mathfrak{m}_j^+\bigr)\bigr\}\bigl\{\bigl[\mathfrak{m}_j^+ - \mathfrak{m}_j^-\bigr] + \sqrt{\varepsilon}\sigma\bigl(\mathfrak{m}_j^+\bigr)\bigr\}.$$

Proposition 5.3. *Assume that $q \geq 1$. Then,*

$$c(\varepsilon) = \bigl[1 + o(\sqrt{\varepsilon})\bigr] Z_\varepsilon e^{H/\varepsilon}.$$

Of course, a sharp asymptotic for the pre-factor F_ε can be derived from Propositions 5.2 and 5.3.

5.3. *Proofs*

We present in this subsection the proofs of Propositions 5.2 and 5.3. We start with an elementary observation.

Lemma 5.4. *The quasi-potential $V(\cdot)$ is continuous.*

Proof. This result follows from Assertion 4.B. On each landscape Λ_n the quasi-potential $\widehat{V}$ differs from S by an additive constant. At the boundary of the landscape Λ_n, $\widehat{V}(\ell_n) = \widehat{V}(\mathfrak{L}^+_{n+1}) = 0$. On each saddle interval Σ_n, $\widehat{V}$ vanishes, which proves the continuity of $\widehat{V}$, and therefore the one of V. □

We continue with a uniform bound for the density π_ε on the landscapes.

Lemma 5.5. *There exists a continuous function $\Xi : \mathbb{R}_+ \to \mathbb{R}_+$ vanishing at the origin such that for each $1 \le n \le \boldsymbol{m}$,*

$$\sup_{x\in\Lambda_n} \frac{1}{\sqrt{\varepsilon}} e^{\widehat{V}(x)/\varepsilon} \big|\pi_\varepsilon(x) - e^{-\widehat{V}(x)/\varepsilon}\big\{G_0(x) + \sqrt{\varepsilon}G_1(x)\big\}\big| \le \Xi(\varepsilon).$$

Proof. Fix a landscape $\Lambda_n = [\ell_n, \mathfrak{L}^+_{n+1}]$ and $x \in \Lambda_n$. Since $S(z(x)) = S(\ell_n)$ on this landscape, rewrite $\pi_\varepsilon(x)$ as

$$e^{[S(\ell_n)-S(x)]/\varepsilon} \int_x^{x+1} e^{[S(y)-S(\ell_n)]/\varepsilon}\, dy = e^{-\widehat{V}(x)/\varepsilon} \int_x^{x+1} e^{[S(y)-S(\ell_n)]/\varepsilon}\, dy.$$

It remains to estimate the integral. Note that $S(y) - S(\ell_n) \le 0$ for $y \ge x$.

The integral is estimated in three steps. Recall from (4.4) the definition of the local maxima $\ell^+_n < \mathfrak{M}^+_{a(n,1)} < \cdots < \mathfrak{M}^+_{a(n,r_n-1)} < \mathfrak{L}^-_{n+1}$ of S such that $S(\mathfrak{M}^+_{a(n,k)}) = S(\ell_n)$. We first consider the integral over the intervals $[\mathfrak{M}^-_{a(n,k)}, \mathfrak{M}^+_{a(n,k)}]$. Then, over the intervals $[\mathfrak{M}^+_{a(n,k)} + \eta, \mathfrak{M}^+_{a(n,k+1)} - \eta]$ for some $\eta > 0$. Finally on the sets $[\mathfrak{M}^+_{a(n,k)}, \mathfrak{M}^+_{a(n,k)} + \eta]$ and $[\mathfrak{M}^-_{a(n,k)} - \eta, \mathfrak{M}^-_{a(n,k)}]$.

Let $\mathcal{N}_n$ be the set of points x' in the landscape Λ_n such that $S(x') = S(\ell_n)$. With the notation just introduced,

$$\mathcal{N}_n = \big\{x' \in \Lambda_n : S\big(x'\big) = S(\ell_n)\big\} = \bigcup_{k=0}^{r_n} \big[\mathfrak{M}^-_{a(n,k)}, \mathfrak{M}^+_{a(n,k)}\big], \tag{5.8}$$

provided $\mathfrak{M}^\pm_{a(n,r_n)} = \mathfrak{L}^\pm_{n+1}$, $\mathfrak{M}^\pm_{a(n,0)} = \ell^\pm_n$, $\ell^-_n = \ell_n$. Of course, some of these intervals may be reduced to points. Since $S(y) = S(\ell_n)$ on $\mathcal{N}_n$,

$$\int_x^{x+1} e^{[S(y)-S(\ell_n)]/\varepsilon}\, dy = \int_x^{x+1} \chi_{\mathcal{N}_n}(y)\, dy + \int_{[x,x+1]\setminus\mathcal{N}_n} e^{[S(y)-S(\ell_n)]/\varepsilon}\, dy.$$

The first term on the right hand side is equal to $G_0(x)$.

We turn to the second integral. We first estimate the integral over open intervals between the maxima. Consider each local maximum $\mathfrak{M}^+_{a(n,k)}$, $1 \le k \le r_n$. Note that the first one, $\mathfrak{M}^+_{a(n,0)} = \ell^+_n$, has not been included and will be treated separately. At each of these points $b(\mathfrak{M}^+_{a(n,k)}) = 0$ and, by assumption (H3), $b'(\mathfrak{M}^+_{a(n,k)}+) > 0$. Choose $\eta > 0$ small enough such that $b'(y) \ge (1/2)b'(\mathfrak{M}^+_{a(n,k)}+)$ for all $y \in [\mathfrak{M}^+_{a(n,k)}, \mathfrak{M}^+_{a(n,k)} + \eta]$ and all k.

Repeat the same procedure for the left endpoints $\mathfrak{M}^-_{a(n,k)}$, $1 \le k \le r_n$. For the point $\mathfrak{M}^+_{a(n,0)} = \ell^+_n$, either $b(\ell^+_n) = 0$ or $b(\ell^+_n) > 0$. In the former case, by assumption (H3), $b'(\ell^+_n+) > 0$, and we may choose $\eta > 0$ small enough such that $b'(y) \ge (1/2)b'(\ell^+_n+)$ for all $y \in [\ell^+_n, \ell^+_n + \eta]$. In the latter case, choose $\eta > 0$ such that $b(y) \ge b(\ell^+_n)/2$ for all $y \in [\ell^+_n, \ell^+_n + \eta]$.

Recall that the landscape $\Lambda_n = [\ell_n, \mathfrak{L}^+_{n+1}] = [\mathfrak{M}^-_{a(n,0)}, \mathfrak{M}^+_{a(n,r_n)}]$. Hence, for any $x \in \Lambda_n$, the interval $[x, x+1]$ is contained in $[\ell_n, \mathfrak{L}^+_{n+1} + 1]$. Let $\mathfrak{C}_n \subset \mathbb{R}$ be the closed set given by

$$\mathfrak{C}_n = \big[\ell^+_n + \eta, \mathfrak{L}^+_{n+1} + 1\big] \setminus \left\{\bigcup_{k=1}^{r_n} \big(\mathfrak{M}^-_{a(n,k)} - \eta, \mathfrak{M}^+_{a(n,k)} + \eta\big)\right\}.$$

For any $y \in \mathfrak{C}_n$, $S(y) < S(\ell_n)$. There exists, therefore, a constant $c(\eta) > 0$ such that $S(y) \le S(\ell_n) - c(\eta)$ for all $y \in \mathfrak{C}_n$. Hence,

$$\int_x^{x+1} \chi_{\mathfrak{C}_n}(y) e^{[S(y)-S(\ell_n)]/\varepsilon}\, dy \le e^{-c(\eta)/\varepsilon}.$$

In view of formula (5.8) for the set $\mathcal{N}_n$, it remains to estimate the integral on the intervals $[\ell_n^+, \ell_n^+ + \eta]$, $[\mathfrak{M}_{a(n,k)}^- - \eta, \mathfrak{M}_{a(n,k)}^-]$, $[\mathfrak{M}_{a(n,k)}^+, \mathfrak{M}_{a(n,k)}^+ + \eta]$. Let

$$\mathcal{D}_n = [\ell_n^+, \ell_n^+ + \eta] \bigcup_{k=1}^{r_n} [\mathfrak{M}_{a(n,k)}^- - \eta, \mathfrak{M}_{a(n,k)}^-] \bigcup_{k=1}^{r_n} [\mathfrak{M}_{a(n,k)}^+, \mathfrak{M}_{a(n,k)}^+ + \eta].$$

Assume that $b(\ell_n^+) = 0$. By Assertions 5.A and 5.B below,

$$\left| \int_{[x,x+1]\cap\mathcal{D}_n} e^{[S(y)-S(\ell_n)]/\varepsilon}\, dy - \sqrt{\varepsilon} G_1(x) \right| \le \sqrt{\varepsilon}\, \Xi(\varepsilon),$$

where

$$\Xi(\varepsilon) = C_0 \Xi\big(b'\big(\ell_n^+ +\big), \varepsilon\big) + C_0 \sum_{k=1}^{r_n} \Xi\big(b'\big(\mathfrak{M}_{a(n,k)}^\pm \pm\big), \varepsilon\big).$$

In the second sum, it has to be understood that there are two sums, one for the terms $\mathfrak{M}_{a(n,k)}^-$ and one for $\mathfrak{M}_{a(n,k)}^+$.

If $b(\ell_n^+) > 0$, $\ell_n = \ell_n^+$, and by the choice of η and Assertion 5.C below,

$$\left| \int_{[x,x+1]\cap[\ell_n^+, \ell_n^+ + \eta]} e^{[S(y)-S(\ell_n)]/\varepsilon}\, dy \right| \le \frac{2\varepsilon}{b(\ell_n)}.$$

This completes the proof of the lemma. □

Proof of Proposition 5.2. Fix $x \in \mathbb{T}$. The case where x belongs to some landscape has been considered in the previous lemma. Consider a saddle interval Σ_n. Recall the definition of the intervals $[c_{n,j}^-, c_{n,j}^+]$, $1 \le j \le s_n$ introduced in (5.5). If $x \in [c_{n,j}^-, c_{n,j}^+]$, since $z(x) = c_{n,j}^+$

$$\pi_\varepsilon(x) = e^{-\widehat{V}(x)/\varepsilon} \int_x^{x+1} e^{[S(y)-S(c_{n,j}^+)]/\varepsilon}\, dy.$$

On the interval $[x, c_{n,j}^+]$, $S(y) = S(c_{n,j}^+)$. Hence the integral on this interval is equal to $c_{n,j}^+ - x = G_0(x)$.

By assumption (H3), $b'(c_{n,j}^+ +) > 0$. Let $\eta > 0$ such that $b'(y) > b'(c_{n,j}^+ +)/2$ for all $y \in [c_{n,j}^+, c_{n,j}^+ + \eta]$. Since $S(y) < S(c_{n,j}^+)$ for all $y > c_{n,j}^+$, there exists $c(\eta) > 0$ such that $S(y) \le S(c_{n,j}^+) - c(\eta)$ for all $y \ge c_{n,j}^+ + \eta$. Hence,

$$\int_{c_{n,j}^+}^{x+1} e^{[S(y)-S(c_{n,j}^+)]/\varepsilon}\, dy = \int_{c_{n,j}^+}^{c_{n,j}^+ + \eta} e^{[S(y)-S(c_{n,j}^+)]/\varepsilon}\, dy + R_\varepsilon,$$

where $R_\varepsilon \le e^{-c(\eta)/\varepsilon}$. By Assertion 5.A, the last integral is equal to $[1 + o(1)]\sqrt{\varepsilon} G_1(x)$. This completes the proof in the case where $x \in [c_{n,j}^-, c_{n,j}^+]$. □

In the case where $x \in \mathcal{G}_n$, the statement of the proposition follows from Assertion 5.D.

Proof of Proposition 5.3. Recall the definition of the set $\mathbb{I}$ introduced in (5.7). Fix $\eta > 0$, to be chosen later, and let $\mathcal{B}_\eta$ be an η-neighborhood of the global minima of V:

$$\mathcal{B}_\eta = \bigcup_{j\in\mathbb{I}} \big(\mathfrak{m}_j^- - \eta, \mathfrak{m}_j^+ + \eta\big).$$

Since, by Lemma 5.4, $\widehat{V}$ is continuous and since $\widehat{V} > -H$ on the closed set $\mathcal{B}_\eta^c$, there exists $c(\eta) > 0$ such that

$$\inf\big\{\widehat{V}(\theta) : \theta \notin \mathcal{B}_\eta\big\} \ge -H + c(\eta). \tag{5.9}$$

Hence, as

$$\sup\{S(y)-S(x):x\le y\le x+1\}=S\big(z(x)\big)-S(x)=-\widehat{V}(x),$$

for every $\eta>0$,

$$\int_{\mathcal{B}_\eta^c}\pi_\varepsilon(x)\,dx=\int_{\mathcal{B}_\eta^c}dx\int_x^{x+1}e^{[S(y)-S(x)]/\varepsilon}\,dy\le e^{[H-c(\eta)]/\varepsilon}.$$

We examine the integral of π_ε on the set $\mathcal{B}_\eta$. Each set $[\mathfrak{m}_j^-,\mathfrak{m}_j^+]$ is contained in the interior of a valley. Choose η small enough for each $[\mathfrak{m}_j^- -\eta,\mathfrak{m}_j^+ +\eta]$ to be contained in the same valley. In this case, by Lemma 5.5, and since G_0 and G_1 are constant in the valleys

$$\int_{\mathfrak{m}_j^- -\eta}^{\mathfrak{m}_j^+ +\eta}\pi_\varepsilon(x)\,dx=\big\{G_0\big(\mathfrak{m}_j^-\big)+\sqrt{\varepsilon}G_1\big(\mathfrak{m}_j^-\big)\pm\sqrt{\varepsilon}\,\Xi(\varepsilon)\big\}\int_{\mathfrak{m}_j^- -\eta}^{\mathfrak{m}_j^+ +\eta}e^{-\widehat{V}(x)/\varepsilon}\,dx.$$

Since $[\mathfrak{m}_j^- -\eta,\mathfrak{m}_j^+ +\eta]$ is contained in a landscape, since in each landscape $\widehat{V}$ and S differ only by an additive constant, and since $\widehat{V}(\mathfrak{m}_j^+)=-H$, on $[\mathfrak{m}_j^- -\eta,\mathfrak{m}_j^+ +\eta]$, $\widehat{V}(x)=\widehat{V}(x)-\widehat{V}(\mathfrak{m}_j^+)-H=S(x)-S(\mathfrak{m}_j^+)-H$. Hence,

$$\int_{\mathfrak{m}_j^- -\eta}^{\mathfrak{m}_j^+ +\eta}e^{-\widehat{V}(x)/\varepsilon}\,dx=e^{H/\varepsilon}\int_{\mathfrak{m}_j^- -\eta}^{\mathfrak{m}_j^+ +\eta}e^{-[S(x)-S(\mathfrak{m}_j^+)]/\varepsilon}\,dx.$$

Choose η small enough to fulfill the assumptions of Assertions 5.A, 5.B (with the obvious modifications since $b'(\mathfrak{m}_j^+)<0$). By these results,

$$\int_{\mathfrak{m}_j^- -\eta}^{\mathfrak{m}_j^+ +\eta}e^{-[S(x)-S(\mathfrak{m}_j^+)]/\varepsilon}\,dx=\big\{\big[\mathfrak{m}_j^+-\mathfrak{m}_j^-\big]+\big[1+o(1)\big]\sqrt{\varepsilon}\sigma\big(\mathfrak{m}_j^+\big)\big\},$$

where $\sigma(\mathfrak{m}_j^+)$ has been introduced in (5.6).

Putting together the previous estimates yields that

$$c(\varepsilon)=\sum_{j\in\mathbb{I}}\big\{G_0\big(\mathfrak{m}_j^+\big)+\big[1+o(1)\big]\sqrt{\varepsilon}G_1\big(\mathfrak{m}_j^+\big)\big\}\big\{\big[\mathfrak{m}_j^+-\mathfrak{m}_j^-\big]+\big[1+o(1)\big]\sqrt{\varepsilon}\sigma\big(\mathfrak{m}_j^+\big)\big\}e^{H/\varepsilon},$$

which completes the proof of the proposition. □

We conclude this section with some estimates used in the proofs above.

Remark 5.6. The proof of these estimates relies on a Taylor expansion of the function S around the local maxima of this function. We need in this argument S'' [that is b'] to be Lipschitz continuous. It is for this reason that we assumed b to be in C^2 in the intervals $[r_j,l_{j+1}]$. We could have assumed the weaker assumption that b' is Lipschitz continuous on these intervals.

Denote by K_0 the Lipschitz continuity constant of b'.

Assertion 5.A. *Let $x\in\mathbb{R}$ be a point such that $b(x)=0$, $b'(x+)>0$. Let $\eta>0$ be such that $b'(y)\ge(1/2)b'(x+)$ for all $y\in[x,x+\eta]$. Then, there exists a finite constant C_0, which depends only on K_0, and a function $\Xi:\mathbb{R}_+^2\to\mathbb{R}_+$ such that $\lim_{\varepsilon\to0}\Xi(a,\varepsilon)=0$ for all $a>0$, and for which*

$$\left|\int_x^{x+\eta}e^{[S(y)-S(x)]/\varepsilon}\,dy-\sqrt{\frac{\pi\varepsilon}{2b'(x+)}}\right|\le C_0\sqrt{\varepsilon}\,\Xi\big(b'(x+),\varepsilon\big).$$

Proof. We derive an upper bound for the integral. The lower bound is obtained by changing $+$ signs into $-$ signs.

Let $\delta = \delta(\varepsilon) > 0$ be a sequence such that $\delta^3 \ll \varepsilon \ll \delta^2$. We first estimate the integral in the interval $[x, x+\delta]$. Since $S''(x+) = -b'(x+) < 0$ and b' is uniformly Lipschitz continuous, in view of the properties of δ, a Taylor expansion and a change of variables yield that

$$\int_x^{x+\delta} e^{[S(y)-S(x)]/\varepsilon}\, dy \le \bigl[1 + C_0\bigl(\delta^3/\varepsilon\bigr)\bigr] \int_0^\delta e^{-(1/2)b'(x+)z^2/\varepsilon}\, dz$$
$$\le \bigl[1 + C_0\bigl(\delta^3/\varepsilon\bigr)\bigr]\sqrt{\frac{\pi\varepsilon}{2b'(x+)}}.$$

It remains to estimate the integral on the interval $[x+\delta, x+\eta]$. By assumption, $S''(y) \le (1/2)S''(x)$ for all $y \in [x, x+\eta]$. Hence,

$$\int_{x+\delta}^{x+\eta} e^{[S(y)-S(x)]/\varepsilon}\, dy \le \int_\delta^\infty e^{S''(x+)y^2/4\varepsilon}\, dy \le \sqrt{\varepsilon}\,\Xi\bigl(b'(x+), \varepsilon\bigr).$$

This proves the assertion. □

The same argument yields the next assertion.

Assertion 5.B. *Let $x \in \mathbb{R}$ be a point such that $b(x) = 0$, $b'(x-) > 0$. Let $\eta > 0$ be such that $b'(y) \ge (1/2)b'(x-)$ for all $y \in [x-\eta, x]$. Then, there exists a finite constant C_0, which depends only on K_0, and a function $\Xi : \mathbb{R}_+^2 \to \mathbb{R}_+$ such that $\lim_{\varepsilon\to 0} \Xi(a, \varepsilon) = 0$ for all $a > 0$, and for which*

$$\left|\int_{x-\eta}^{x} e^{[S(y)-S(x)]/\varepsilon}\, dy - \sqrt{\frac{\pi\varepsilon}{2b'(x-)}}\right| \le C_0\sqrt{\varepsilon}\,\Xi\bigl(b'(x-), \varepsilon\bigr).$$

It remains to consider the case where $b(x) > 0$.

Assertion 5.C. *Let $x \in \mathbb{R}$ be a point such that $b(x) > 0$. Let $\eta > 0$ such that $b(y) \ge b(x)/2$ for all $y \in [x, x+\eta]$. Then,*

$$\int_x^{x+\eta} e^{[S(y)-S(x)]/\varepsilon}\, dy \le \frac{2\varepsilon}{b(x)}.$$

Proof. By a Taylor expansion and by hypothesis,

$$\int_x^{x+\eta} e^{[S(y)-S(x)]/\varepsilon}\, dy \le \int_x^{x+\eta} e^{-b(x)(y-x)/2\varepsilon}\, dy \le \int_0^\infty e^{-b(x)z/2\varepsilon}\, dz = \frac{2\varepsilon}{b(x)}.$$ □

If we assume that $S(y) < S(x)$ for all $x < y$, we may estimate the integral over the interval $[x, x+1]$.

Assertion 5.D. *Let $x \in \mathbb{R}$ be a point such that $b(x) > 0$. Assume, furthermore, that $S(y) < S(x)$ for all $y \in (x, x+1]$. Then,*

$$\int_x^{x+1} e^{[S(y)-S(x)]/\varepsilon}\, dy = \bigl[1 + o(1)\bigr]\frac{\varepsilon}{b(x)}.$$

Proof. Let $\delta = \delta(\varepsilon) > 0$ be a sequence such that $\delta^2 \ll \varepsilon \ll \delta$. By the Taylor expansion and an elementary computation, as $S'(x) = -b(x)$,

$$\int_x^{x+\delta} e^{[S(y)-S(x)]/\varepsilon}\, dy = \int_0^\delta e^{[-b(x)y+O(\delta^2)]/\varepsilon}\, dy = \bigl[1 + o(1)\bigr]\frac{\varepsilon}{b(x)}.$$

Let $\alpha = b(x)/2 > 0$. There exists $\eta > 0$ such that $S(y) - S(x) \le -\alpha(y-x)$ for all $y \in [x, x+\eta]$. On the other hand, since $S(y) < S(x)$ for all $y \in [x+\eta, x+1]$ and since S is continuous, there exists $\kappa > 0$ such that $S(y) \le S(x) - \kappa$ for all $y \in [x+\eta, x+1]$. Therefore,

$$\int_{x+\delta}^{x+\eta} e^{[S(y)-S(x)]/\varepsilon}\,dy \le \int_{\delta}^{\eta} e^{-(\alpha/\varepsilon)y}\,dy = o(1)\varepsilon,$$

$$\int_{x+\eta}^{x+1} e^{[S(y)-S(x)]/\varepsilon}\,dy \le e^{-\kappa/\varepsilon} = o(1)\varepsilon.$$

The assertion follows from the three previous estimates. □

5.4. *When the set* $\{\theta \in \mathbb{T} : b(\theta) = 0\}$ *is finite*

We present in this subsection a formula for the pre-factor in the case where the connected components of the set $\{\theta \in \mathbb{T} : b(\theta) = 0\}$ are points.

(H4) Assume that the connected components of the set $\{\theta \in \mathbb{T} : b(\theta) = 0\}$ are points, that b is of class $C^2(\mathbb{T})$ and that $b'(\theta) \neq 0$ for all $\theta \in \mathbb{T}$ such that $b(\theta) = 0$ [that is $S''(x) \neq 0$ at the critical points of S].

Note that these assumptions imply that $\ell_n^+ = \ell_n$, $b(\ell_n) > 0$ for all left endpoints of a landscape and that $\mathfrak{M}_k^+ = \mathfrak{M}_k^-$, $\mathfrak{m}_k^+ = \mathfrak{m}_k^-$ for all k. Moreover, the sets $\mathcal{F}_n$ introduced in (5.5) are empty, so that $\Sigma_n = \mathcal{G}_n$.

Set $\mathfrak{M}_k := \mathfrak{M}_k^+$, $\mathfrak{m}_k := \mathfrak{m}_k^+$, $\mathfrak{L}_k := \mathfrak{L}_k^+$ for all indices k. Fix a landscape Λ_n. Under the previous hypotheses, $G_0 \equiv 0$ and G_1 is given by (5.4). In a saddle interval Σ_n, $G_0 \equiv 0$ and $G_1 \equiv 0$, while the function G_2 is unchanged. The weights $\omega(\mathfrak{M}_k)$, $\sigma(\mathfrak{m}_k)$, $1 \le k \le q$, become

$$\omega(\mathfrak{M}_k) = \sqrt{\frac{2\pi}{b'(\mathfrak{M}_k)}}, \qquad \sigma(\mathfrak{m}_k) = \sqrt{\frac{2\pi}{-b'(\mathfrak{m}_k)}}.$$

Set

$$Z = \sum_{j \in \mathbb{I}} G_1(\mathfrak{m}_j)\sigma(\mathfrak{m}_j).$$

By the definition of Z_ε and by Proposition 5.3, $Z_\varepsilon = \varepsilon Z$ and

$$c(\varepsilon) = \big[1 + o(\sqrt{\varepsilon})\big] Z\varepsilon e^{H/\varepsilon}. \tag{5.10}$$

Thus, Proposition 5.2 can be restated in this context as follows.

Proposition 5.7. *Assume that hypotheses* (H4) *are in force. Then*,

(1) (*Pre-factor on the landscapes*)

$$\lim_{\varepsilon \to 0} \sup_{x \in \Lambda} \sqrt{\varepsilon} e^{V(x)/\varepsilon} \left| m_\varepsilon(x) - \frac{1}{Z}\frac{1}{\sqrt{\varepsilon}} G_1(x) e^{-V(x)/\varepsilon} \right| = 0.$$

(2) (*Pre-factor on the saddle intervals*)

$$m_\varepsilon(x) = \big[1 + o(1)\big] \frac{1}{Z} G_2(x) e^{-V(x)/\varepsilon}, \quad x \in \Sigma.$$

Remark 5.8. The results of this article remain in force if we add a $(d-1)$-transversal drift. More precisely, consider the diffusion on $\mathbb{T}^d$ given by

$$dX_\varepsilon(t) = \boldsymbol{b}\big(X_\varepsilon(t)\big)\,dt + \sqrt{2\varepsilon}\,dW_t,$$

where W_t is a Brownian motion on $\mathbb{T}^d$, and $\boldsymbol{b} = (b_1, \dots, b_d) : \mathbb{T}^d \to \mathbb{R}$ a drift. The same results hold provided that

$$b_1(x_1, \dots, x_d) = b_1(x_1) \quad \text{and} \quad \sum_{j=2}^{d} (\partial_{x_j} b_j)(x) = 0.$$

5.5. *The Hamilton–Jacobi equation*

We examine in this subsection the asymptotic behavior, as $\varepsilon \to 0$, of the solution of the Hamilton–Jacobi equation satisfied by the pre-factor of the stationary measure. We consider this problem under the assumptions (H4).

Since m_ε is the density of the stationary state,

$$\varepsilon m_\varepsilon'' - (b m_\varepsilon)' = 0. \tag{5.11}$$

Since the quasi-potential V is not continuously differentiable, but only smooth by parts, we consider the previous equation separately on the landscapes Λ_n and on the saddle intervals Σ_n, $1 \le n \le \boldsymbol{m}$.

Inserting expression (5.2) for the stationary state m_ε in (5.11) yields the following equation:

$$\varepsilon F_\varepsilon'' + F_\varepsilon' b = 0 \quad \text{on } \Lambda \quad \text{and} \quad \varepsilon F_\varepsilon'' - F_\varepsilon' b - F_\varepsilon b' = 0 \quad \text{on } \Sigma, \tag{5.12}$$

which is the Hamilton–Jacobi equation for the pre-factor.

Denote by F the solution of the Hamilton–Jacobi equation (5.12) with $\varepsilon = 0$. Clearly, there exist constants c_0 and c_1 such that

$$\begin{aligned} &F(\theta) = c_1 \quad \text{at each connected component of } \{\theta : b(\theta) \neq 0\} \cap \Lambda, \\ &F(\theta) = \frac{c_0}{b(\theta)} \quad \text{at each connected component of } \Sigma. \end{aligned} \tag{5.13}$$

Note that the constants may differ on distinct connected components.

We now compare (5.13) with the asymptotic behavior, as $\varepsilon \to 0$, of the solution of the Hamilton–Jacobi equation on the set Λ. The solution is given by

$$F_\varepsilon(\theta) = c_0 + c_1 \int_{\theta_0}^{\theta} e^{S(y)/\varepsilon} \, dy,$$

for $c_0, c_1 \in \mathbb{R}$, $\theta_0 \in \mathbb{T}$.

Recall from (4.3) that the connected component Λ_n of Λ are intervals of the form $(\ell_n, \mathfrak{L}_{n+1})$, $1 \le n \le \boldsymbol{m}$. Keep in mind that $\mathfrak{L}_{n+1}$ is a local maximum of S and ℓ_n a point such that $S'(\ell_n) < 0$. Moreover, $S(\ell_n) = S(\mathfrak{L}_{n+1})$ and $S(\theta) \le S(\ell_n)$ for all $\theta \in (\ell_n, \mathfrak{L}_{n+1})$.

For $F_\varepsilon(\theta)$ to converge at $\theta = \mathfrak{L}_{n+1}$ to a non trivial value, we have to choose c_1 as $c_1' \varepsilon^{-1/2} \exp\{-S(\ell_n)/\varepsilon\}$ for some $c_1' \in \mathbb{R}$. In contrast, the choice of θ_0 is not important. With this choice,

$$F_\varepsilon(\theta) = c_0 + c_1' \frac{1}{\sqrt{\varepsilon}} \int_{\theta_0}^{\theta} e^{[S(y) - S(\ell_n)]/\varepsilon} \, dy. \tag{5.14}$$

The next result follows from the calculations presented in Assertions 5.A–5.D.

Assertion 5.E. *Fix $\theta_0 \in (\ell_n, \mathfrak{L}_{n+1})$ and consider F_ε given by (5.14). Then, for all $\theta \in [\ell_n, \mathfrak{L}_{n+1}]$,*

$$\boldsymbol{F}(\theta) := \lim_{\varepsilon \to 0} F_\varepsilon(\theta) = c_0 + c_1' \big[G_1(\theta) - G_1(\theta_0) \big].$$

The function $\boldsymbol{F}$ inherits the properties of G_1, it is constant in the valleys $\mathcal{V}_{n,j}$, $1 \le j \le r_n$, and discontinuous at the local maxima $\mathfrak{M}^+_{a(n,j)}$, unless $c_1' = 0$. In particular, it fulfills the conditions in the first line of (5.13).

We set the value of c_1 for $F_\varepsilon(\mathfrak{L}_{n+1})$ to converge. Choosing c_1 for $F_\varepsilon(\theta_1)$ to converge, for some $\theta_1 \in \Lambda_n$ such that $S(\theta_1) < S(\ell_n)$, would produce a limit equal to $\pm\infty$ at every point y such that $S(y) > S(\theta_1)$.

We turn to the set Σ. Fix a connected component $\Sigma_n = (\mathfrak{L}_n, \ell_n)$. An elementary computation yields that the solution of equation (5.12) on Σ_n is given by

$$F_\varepsilon(\theta) = \frac{1}{\varepsilon}\left\{c_0 \int_{\theta_0}^{\theta} e^{[S(y)-S(\theta)]/\varepsilon}\, dy + c_1 e^{-S(\theta)/\varepsilon}\right\} \tag{5.15}$$

for constants $c_0, c_1 \in \mathbb{R}$, which may depend on ε, and some $\theta_0 \in \Sigma_n$ which may also depend on ε.

Assertion 5.F. *There are no choices of the constants $c_0(\varepsilon), c_1(\varepsilon), \theta_0(\varepsilon)$ for which F_ε has a non-trivial limit as $\varepsilon \to 0$.*

Proof. If we set $\theta_0 = \mathfrak{L}_n$, a Taylor expansion yields that (5.15) is equal to

$$F_\varepsilon(\theta) = \frac{1}{\varepsilon}\left\{\big[1+o(1)\big]c_0\sqrt{\frac{\varepsilon\pi}{2|S''(\mathfrak{L}_n)|}}e^{S(\mathfrak{L}_n)/\varepsilon} + c_1\right\}e^{-S(\theta)/\varepsilon}.$$

The expression inside braces is a function of ε which can compensate the factor ε^{-1} or which can be of a smaller order. In any case, this constant is multiplied by $\exp\{-S(\theta)/\varepsilon\}$ which may converges for one specific $\theta \in \Sigma_n$ but which will diverge for all other θ. Hence, if $\theta_0 = \mathfrak{L}_n$ there is no choice of $c_0(\varepsilon)$, $c_1(\varepsilon)$ which provide a non-trivial limit for (5.15). A similar analysis can be carried through if θ_0 is chosen in $(\mathfrak{L}_n, \ell_n]$, which proves the assertion. □

The previous assertion shows that on the set Σ the solution $F_\varepsilon(\theta)$ of (5.12) does not converge, as $\varepsilon \to 0$, to the solution $F(\theta)$ of (5.13) unless we consider the trivial solutions $F_\varepsilon(\theta) = F(\theta) = 0$.

6. Equilibrium potential and capacities

We estimate in this section capacities between wells. We start with an explicit formula for the adjoint of L_ε in $L^2(\mu_\varepsilon)$, the Hilbert space of measurable functions $f : \mathbb{T} \to \mathbb{R}$ endowed with the scalar product given by

$$\langle f, g\rangle_{\mu_\varepsilon} = \int_{\mathbb{T}} f(\theta)g(\theta)\mu_\varepsilon(d\theta).$$

Integrating the equation (5.11) once provides that

$$m'_\varepsilon(\theta) = -R_\varepsilon + \frac{1}{\varepsilon}b(\theta)m_\varepsilon(\theta), \quad \text{where } R_\varepsilon = \frac{1}{c(\varepsilon)}\big(1 - e^{-B/\varepsilon}\big). \tag{6.1}$$

Note that R_ε is positive and that it vanishes if $B = 0$.

Denote by L_ε^* the adjoint operator of L_ε in $L^2(\mu_\varepsilon)$. It follows from (6.1) that for every twice continuously differentiable function $f : \mathbb{T} \to \mathbb{R}$,

$$\big(L_\varepsilon^* f\big)(\theta) = \left(b(\theta) - \frac{2\varepsilon R_\varepsilon}{m_\varepsilon(\theta)}\right)f'(\theta) + \varepsilon f''(\theta).$$

In particular, $L_\varepsilon^* = L_\varepsilon$ if $B = 0$, and the symmetric part of the generator, denoted by $S_\varepsilon = (1/2)(L_\varepsilon + L_\varepsilon^*)$, is given by

$$(S_\varepsilon f)(\theta) = \left(b(\theta) - \frac{\varepsilon R_\varepsilon}{m_\varepsilon(\theta)}\right)f'(\theta) + \varepsilon f''(\theta).$$

The Dirichlet form, denoted by $D_\varepsilon(\cdot)$, associated to the generator L_ε is given by

$$D_\varepsilon(f) := -\int_{\mathbb{T}} f(\theta)(S_\varepsilon f)(\theta)m_\varepsilon(\theta)\, d\theta = \varepsilon\int_{\mathbb{T}}\big[f'(\theta)\big]^2 m_\varepsilon(\theta)\, d\theta. \tag{6.2}$$

Equilibrium potential and capacity

Fix two disjoint closed intervals $\mathcal{A}_1 = [\theta_1^-, \theta_1^+]$, $\mathcal{A}_2 = [\theta_2^-, \theta_2^+]$ of $\mathbb{T}$. Without loss of generality, we suppose that

$$0 \le \theta_1^- \le \theta_1^+ < \theta_2^- \le \theta_2^+ < 1. \tag{6.3}$$

Note that we allow the intervals to be reduced to a point. The unique solution to the elliptic problem

$$\begin{cases} (L_\varepsilon f)(\theta) = 0, & \theta \in \mathbb{T} \setminus (\mathcal{A}_1 \cup \mathcal{A}_2), \\ f(\theta) = \chi_{\mathcal{A}_1}(\theta), & \theta \in \mathcal{A}_1 \cup \mathcal{A}_2 \end{cases}$$

is called the equilibrium potential between the sets $\mathcal{A}_1$ and $\mathcal{A}_2$, and is denoted by $h_{\mathcal{A}_1,\mathcal{A}_2} = h^\varepsilon_{\mathcal{A}_1,\mathcal{A}_2}$.

In dimension 1, an explicit formula for the equilibrium potential is available, a straightforward computation shows that

$$h_{\mathcal{A}_1,\mathcal{A}_2}(\theta) = \frac{\int_\theta^{\theta_2^-} e^{S(y)/\varepsilon}\, dy}{\int_{\theta_1^+}^{\theta_2^-} e^{S(y)/\varepsilon}\, dy} \chi_{[\theta_1^+,\theta_2^-]}(\theta) + \frac{\int_{\theta_2^+}^{\theta} e^{S(y)/\varepsilon}\, dy}{\int_{\theta_2^+}^{1+\theta_1^-} e^{S(y)/\varepsilon}\, dy} \chi_{[\theta_2^+,1+\theta_1^-]}(\theta). \tag{6.4}$$

Define the capacity between $\mathcal{A}_1$ and $\mathcal{A}_2$ as the Dirichlet form of the equilibrium potential:

$$\operatorname{cap}_\varepsilon(\mathcal{A}_1, \mathcal{A}_2) := D_\varepsilon(h_{\mathcal{A}_1,\mathcal{A}_2}) = \varepsilon \int_{\mathbb{T}} \big[h'_{\mathcal{A}_1,\mathcal{A}_2}(\theta)\big]^2 m_\varepsilon(\theta)\, d\theta.$$

We show in Assertion 6.B below that

$$\operatorname{cap}_\varepsilon(\mathcal{A}_1, \mathcal{A}_2) = \varepsilon\big\{h'_{\mathcal{A}_1,\mathcal{A}_2}\big(\theta_1^-\big)m_\varepsilon\big(\theta_1^-\big) - h'_{\mathcal{A}_1,\mathcal{A}_2}\big(\theta_1^+\big)m_\varepsilon\big(\theta_1^+\big)\big\}. \tag{6.5}$$

Moreover, as $h_{\mathcal{A}_2,\mathcal{A}_1} = 1 - h_{\mathcal{A}_1,\mathcal{A}_2}$,

$$\operatorname{cap}_\varepsilon(\mathcal{A}_2, \mathcal{A}_1) = \operatorname{cap}_\varepsilon(\mathcal{A}_1, \mathcal{A}_2).$$

Estimation of capacity

We present in Propositions 6.1–6.3 below sharp estimates of the capacity between two sets which satisfy the conditions below.

Assume that the intervals $\mathcal{A}_1 = [\theta_1^-, \theta_1^+]$, $\mathcal{A}_2 = [\theta_2^-, \theta_2^+]$ represent wells (cf. Section 7.1) in the sense that

$$V(\theta) < V\big(\theta_i^-\big) = V\big(\theta_i^+\big) < H \quad \text{for all } \theta \in \big(\theta_i^-, \theta_i^+\big), i = 1, 2. \tag{6.6}$$

We refer to Figure 3. In particular, each interval $\mathcal{A}_i$ is contained in some valley, denoted by $\mathcal{W}_i = \mathcal{V}_{n(i),k(i)}$, of some landscape $\Lambda'_i = \Lambda_{n(i)}$. Of course, the valleys and the landscapes may coincide or not.

As the sets $\mathcal{A}_i$ are contained in valleys and the pre-factors G_a, $a = 0, 1$, are constant in valleys,

$$G_a\big(\theta_i^-\big) = G_a\big(\theta_i^+\big), \quad a = 0, 1, i = 1, 2. \tag{6.7}$$

This identity will be used repeatedly below to replace θ_i^- by θ_i^+.

By Assertion 4.B, V and S differ only by an additive constant on the valley $\mathcal{W}_i$. In particular,

$$S(\theta) = V(\theta) + C_i, \quad \theta \in \mathcal{A}_i,$$

and V is differentiable in $\mathcal{W}_i$. It follows from (6.6) that $V'(\theta_i^-) \le 0 \le V'(\theta_i^+)$. We assume a strict inequality:

$$V'\big(\theta_i^-\big) < 0 < V'\big(\theta_i^+\big). \tag{6.8}$$

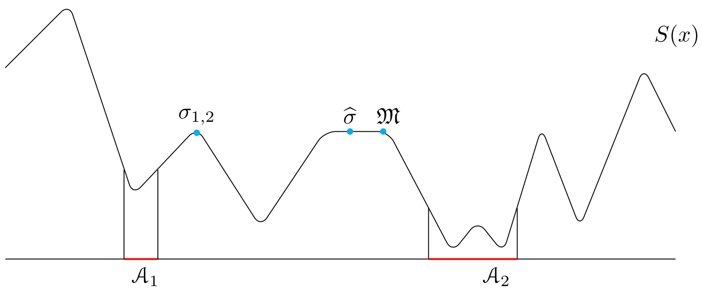

Fig. 3. This figure represents two disjoint intervals $\mathcal{A}_1$, $\mathcal{A}_2$ of $\mathbb{T}$ which belong to a valley $\mathcal{W}$. In this case, the energy barrier between $\mathcal{A}_1$ and $\mathcal{A}_2$ is much smaller inside the valley (that is, in the interval $[\theta_1^+, \theta_2^-]$) than outside it. The calculation of the capacity is thus reduced to a computation in the latter interval.

Two points σ, resp. σ^* in (θ_1^+, θ_2^-), resp. $(\theta_2^+, 1+\theta_1^-)$, are called saddle points between $\mathcal{A}_1$ and $\mathcal{A}_2$ if

$$S(\sigma) = \max\{S(y) : \theta_1^+ \le y \le \theta_2^-\},$$
$$S(\sigma^*) = \max\{S(y) : \theta_2^+ \le x \le 1+\theta_1^-\}.$$

Of course, there may be more than one, but let us fix two saddle points between $\mathcal{A}_1$ and $\mathcal{A}_2$, $\sigma_{1,2} \in (\theta_1^+, \theta_2^-)$, $\sigma_{2,1} \in (\theta_2^+, 1+\theta_1^-)$.

Observe that $V(\sigma_{2,1}) = H$ if $V(\sigma_{1,2}) < H$. Indeed, if $V(\sigma_{1,2}) < H$, $\mathcal{A}_1$, $\mathcal{A}_2$ belong to the same valley. This implies that $V(\sigma) = H$ for all saddle points in $(\theta_2^+, 1+\theta_1^-)$.

In the computation of the capacity between $\mathcal{A}_1$, $\mathcal{A}_2$, three cases emerge. The sets $\mathcal{A}_1$, $\mathcal{A}_2$ may belong to the same valley, to different valleys but to the same landscape, or to different landscapes. Consider first the case, illustrated in Figure 3, in which both sets belong to the same valley.

Assume that the sets $\mathcal{A}_i$ are contained in a valley $\mathcal{W} = (\mathfrak{w}^-, \mathfrak{w}^+)$. If $\mathfrak{w}^- < \theta_1^- < \theta_2^+ < \mathfrak{w}^+$, let $E_{1,2}$ be the set of local maxima $\mathfrak{M}_k^+$ of S in (θ_1^+, θ_2^-) such that $S(\mathfrak{M}_k^+) = S(\sigma_{1,2})$:

$$E_{1,2} = \{k \in \{1, \dots, q\} : \mathfrak{M}_k^+ \in [\theta_1^+, \theta_2^-], S(\mathfrak{M}_k^+) = S(\sigma_{1,2})\}. \tag{6.9}$$

If $\mathfrak{w}^- < \theta_2^- < 1+\theta_1^+ < \mathfrak{w}^+$, $E_{1,2}$ represents the set of local maxima $\mathfrak{M}_k^+$ of S in $(\theta_2^+, 1+\theta_1^-)$ such that $S(\mathfrak{M}_k^+) = S(\sigma_{2,1})$. In Figure 3, $E_{1,2} = \{a, b\}$ if $\sigma_{1,2} = \mathfrak{M}_a^+$, $\mathfrak{M} = \mathfrak{M}_b^+$.

For $1 \le i \le q$, let

$$\omega(\mathfrak{M}_i^+) = \omega_+(\mathfrak{M}_i^+) + \omega_-(\mathfrak{M}_i^-). \tag{6.10}$$

Proposition 6.1. *Let $\mathcal{A}_1 = [\theta_1^-, \theta_1^+]$, $\mathcal{A}_2 = [\theta_2^-, \theta_2^+]$ be two intervals satisfying conditions* (6.3), (6.6), (6.8). *Suppose that the sets $\mathcal{A}_1$, $\mathcal{A}_2$ belong to a valley $\mathcal{W} = (\mathfrak{w}^-, \mathfrak{w}^+)$, $1 \le j \le n$. Then,*

$$\mathrm{cap}_\varepsilon(\mathcal{A}_1, \mathcal{A}_2) = [1 + o(\sqrt{\varepsilon})] \frac{\varepsilon}{Z_\varepsilon} \frac{G_0(\theta_1^+) + \sqrt{\varepsilon} G_1(\theta_1^+) + o(\sqrt{\varepsilon})}{\sum_{k \in E_{1,2}} [\mathfrak{M}_k^+ - \mathfrak{M}_k^- + \sqrt{\varepsilon}\omega(\mathfrak{M}_k^+)] + o(\sqrt{\varepsilon})} e^{-V(\sigma_{1,2})/\varepsilon}.$$

We may replace on the right hand side θ_1^+ by θ_2^+ because G_0, G_1 are constant in the valleys.

We turn to the case in which the sets $\mathcal{A}_1$, $\mathcal{A}_2$ belong to different landscapes, so that $z(\theta_1^+) < z(\theta_2^+) < z(1+\theta_1^+)$. Figure 4 illustrates this situation.

Proposition 6.2. *Let $\mathcal{A}_1 = [\theta_1^-, \theta_1^+]$, $\mathcal{A}_2 = [\theta_2^-, \theta_2^+]$ be two intervals satisfying conditions* (6.3), (6.6), (6.8). *Suppose that they belong to different landscapes. Then,*

$$\mathrm{cap}_\varepsilon(\mathcal{A}_1, \mathcal{A}_2) = [1 + o(1)] \frac{\varepsilon}{Z_\varepsilon} e^{-H/\varepsilon}.$$

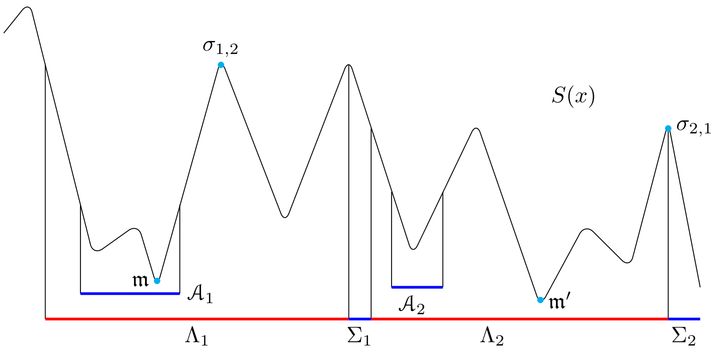

Fig. 4. This figure illustrates the case in which the intervals $\mathcal{A}_i$ belong to different landscapes. Starting from $\mathcal{A}_1$ the process reaches $\mathcal{A}_2$ by surmounting the energetic barrier $[S(\sigma_{1,2}) - S(\mathfrak{m})]/\varepsilon$, while it reaches $\mathcal{A}_1$ when starting from $\mathcal{A}_2$, by surpassing the energetic barrier $[S(\sigma_{2,1}) - S(\mathfrak{m}')]/\varepsilon$ because at the large deviations level, it never visits a landscape Λ_n once in Λ_{n+1}. The capacity represents the height of the saddle point which in this case is equal to H in both cases since $V(\sigma_{1,2}) = V(\sigma_{2,1}) = H$.

For $a = 0, 1$, let

$$(\Delta_{1,2}G_a) = G_a(\theta_1^+) - G_a(\theta_2^+), \qquad (\Delta_{2,1}G_a) = G_a(\theta_2^+) - G_a(\theta_1^+).$$

Suppose that θ_1^+, θ_2^+ belong to the same landscape but to different valleys. Then, either $z(\theta_2^+) = z(\theta_1^+)$ or $z(\theta_2^+) = z(1+\theta_1^+)$. If $z(\theta_2^+) = z(\theta_1^+)$, $G_1(\theta_2^+) < G_1(\theta_1^+)$ so that $(\Delta_{1,2}G_1) > 0$. While, if $z(\theta_2^+) = z(1+\theta_1^+)$, $G_1(\theta_1^+) = G_1(1+\theta_1^+) < G_1(\theta_2^+)$ so that $(\Delta_{2,1}G_1) > 0$. Similar conclusions hold if we replace G_1 by G_0 and a strict inequality by an inequality.

Proposition 6.3. *Let $\mathcal{A}_1 = [\theta_1^-, \theta_1^+]$, $\mathcal{A}_2 = [\theta_2^-, \theta_2^+]$ be two intervals satisfying conditions (6.3), (6.6), (6.8). Suppose that they belong to different valleys, but to the same landscape. Then, if $z(\theta_1^+) = z(\theta_2^+)$,*

$$\mathrm{cap}_\varepsilon(\mathcal{A}_1, \mathcal{A}_2) = \big[1 + o(\sqrt{\varepsilon})\big]\frac{\varepsilon}{Z_\varepsilon} e^{-H/\varepsilon} \frac{G_0(\theta_1^+) + \sqrt{\varepsilon} G_1(\theta_1^+) + o(\sqrt{\varepsilon})}{(\Delta_{1,2}G_0) + \sqrt{\varepsilon}(\Delta_{1,2}G_1) + o(\sqrt{\varepsilon})}.$$

If $z(\theta_2^+) = z(1+\theta_1^+)$, $\mathrm{cap}_\varepsilon(\mathcal{A}_1, \mathcal{A}_2)$ is equal to

$$\big[1 + o(\sqrt{\varepsilon}))\big]\frac{\varepsilon}{Z_\varepsilon} e^{-H/\varepsilon}\bigg(1 + o(1) + \frac{G_0(\theta_1^+) + \sqrt{\varepsilon} G_1(\theta_1^+) + o(\sqrt{\varepsilon})}{(\Delta_{2,1}G_0) + \sqrt{\varepsilon}(\Delta_{2,1}G_1) + o(\sqrt{\varepsilon})}\bigg).$$

The proofs of the previous results rely on the next claim.

Assertion 6.A. *Let $\mathcal{A}_1 = [\theta_1^-, \theta_1^+]$, $\mathcal{A}_2 = [\theta_2^-, \theta_2^+]$ be two intervals satisfying conditions (6.3), (6.6), (6.8). Then,*

$$\begin{aligned}\mathrm{cap}_\varepsilon(\mathcal{A}_1, \mathcal{A}_2) = {} & \big[1 + o(\sqrt{\varepsilon})\big]\frac{\varepsilon}{Z_\varepsilon}\big\{G_0(\theta_1^+) + \sqrt{\varepsilon} G_1(\theta_1^+) + o(\sqrt{\varepsilon})\big\} e^{-V(\theta_1^+)/\varepsilon} \\ & \times \bigg\{\frac{1}{\int_{\theta_2^+}^{1+\theta_1^-} e^{[S(y) - S(1+\theta_1^-)]/\varepsilon}\, dy} + \frac{1}{\int_{\theta_1^+}^{\theta_2^-} e^{[S(y) - S(\theta_1^+)]/\varepsilon}\, dy}\bigg\}.\end{aligned}$$

Proof. Fix two intervals $\mathcal{A}_1 = [\theta_1^-, \theta_1^+]$, $\mathcal{A}_2 = [\theta_2^-, \theta_2^+]$ satisfying the hypotheses of the assertion.

In view of equations (6.4), (6.5),

$$\mathrm{cap}_\varepsilon(\mathcal{A}_1, \mathcal{A}_2) = \varepsilon\bigg\{\frac{e^{S(1+\theta_1^-)/\varepsilon}}{\int_{\theta_2^+}^{1+\theta_1^-} e^{S(y)/\varepsilon}\, dy} m_\varepsilon(\theta_1^-) + \frac{e^{S(\theta_1^+)/\varepsilon}}{\int_{\theta_1^+}^{\theta_2^-} e^{S(y)/\varepsilon}\, dy} m_\varepsilon(\theta_1^+)\bigg\}.$$

Note that in the first ratio on the right hand side we have $1+\theta_1^-$ instead of θ_1^- because we are now working on the line so that the harmonic function is defined on the interval $[\theta_2^+, 1+\theta_1^-]$. As m_ε is periodic, $m_\varepsilon(1+\theta_1^-) = m_\varepsilon(\theta_1^-)$.

Propositions 5.2, 5.3 provide a formula for $m_\varepsilon = \pi_\varepsilon / c(\varepsilon)$. By (6.7), we may replace $G_a(\theta_1^-)$ by $G_a(\theta_1^+)$, $a = 1, 2$. To complete the proof, it remains to recall that $V(\theta_1^-) = V(\theta_1^+)$ by hypothesis. $\square$

Proof of Proposition 6.1. Fix two intervals $\mathcal{A}_1 = [\theta_1^-, \theta_1^+]$, $\mathcal{A}_2 = [\theta_2^-, \theta_2^+]$ satisfying the hypotheses of the proposition.

We estimate the second term inside braces in the formula for the capacity appearing in Assertion 6.A. Since $\mathcal{A}_1$ and $\mathcal{A}_2$ belong to the same valley, by Assertion 4.B, in the interval $[\theta_1^+, \theta_2^-]$, the functions V and S differ only by an additive constant. Recall the definition (6.9) of the set $E_{1,2}$. By the computation performed in Assertion 5.A,

$$\int_{\theta_1^+}^{\theta_2^-} e^{[S(y)-S(\theta_1^+)]/\varepsilon}\, dy = \biggl(\sum_{k \in E_{1,2}} \bigl[\mathfrak{M}_k^+ - \mathfrak{M}_k^-\bigr] + \sqrt{\varepsilon} \sum_{k \in E_{1,2}} \omega\bigl(\mathfrak{M}_k^+\bigr) + o(\sqrt{\varepsilon})\biggr) e^{[S(\sigma_{1,2})-S(\theta_1^+)]/\varepsilon}.$$

Since V and S differ by a constant, we may replace in the previous formula $S(\sigma_{1,2}) - S(\theta_1^+)$ by $V(\sigma_{1,2}) - V(\theta_1^+)$.

We turn to the first term inside braces. As $\mathcal{A}_1$, $\mathcal{A}_2$ belong to the same valley, $z(\theta_2^+) = z(\theta_1^-)$, so that $z(\theta_2^+) < 1+\theta_1^-$ because $z(\theta_1^-) < 1+\theta_1^-$. Therefore, since $S(1+\theta_1^-) = -B + S(\theta_1^-)$,

$$\int_{\theta_2^+}^{1+\theta_1^-} e^{[S(y)-S(1+\theta_1^-)]/\varepsilon}\, dy = e^{[S(z(\theta_2^+))-S(\theta_1^-)+B]/\varepsilon} \int_{\theta_2^+}^{1+\theta_1^-} e^{[S(y)-S(z(\theta_2^+))]/\varepsilon}\, dy.$$

As $z(\theta_2^+) = z(\theta_1^-)$, $S(z(\theta_2^+)) - S(\theta_1^-) = -\widehat{V}(\theta_1^-)$. On the other hand, since $z(\theta_2^+) < 1+\theta_1^-$, by the computation performed in Assertion 5.A

$$\int_{\theta_2^+}^{1+\theta_1^-} e^{[S(y)-S(z(\theta_2^+))]/\varepsilon}\, dy \ge C_0 \sqrt{\varepsilon}$$

for some positive constant C_0 independent of ε. Thus, the right hand side of the penultimate displayed equation is bounded below by $C_0\sqrt{\varepsilon}\exp\{-[\widehat{V}(\theta_1^-) - B]/\varepsilon\}$.

To complete the proof of the proposition, it remains to show that $\widehat{V}(\sigma_{1,2}) < B$, but this is clear because $\widehat{V}(\sigma_{1,2}) \le 0 < B$. $\square$

Proof of Proposition 6.2. In the formula for the capacity of Assertion 6.A, write the second term in the expression inside braces as

$$\int_{\theta_1^+}^{\theta_2^-} e^{[S(y)-S(\theta_1^+)]/\varepsilon}\, dy = e^{[S(z(\theta_1^+))-S(\theta_1^+)]/\varepsilon} \int_{\theta_1^+}^{\theta_2^-} e^{[S(y)-S(z(\theta_1^+))]/\varepsilon}\, dy.$$

Since $z(\theta_1^+) < \theta_2^-$, the integral on the right hand side is equal to $G_0(\theta_1^+) + \sqrt{\varepsilon} G_1(\theta_1^+) + o(\sqrt{\varepsilon})$, while the term appearing in the exponential is equal to $-\widehat{V}(\theta_1^+)$.

We turn to the first term in the expression inside braces in Assertion 6.A. It can be written as

$$e^{[S(z(\theta_2^+))-S(1+\theta_1^-)]/\varepsilon} \int_{\theta_2^+}^{1+\theta_1^-} e^{[S(y)-S(z(\theta_2^+))]/\varepsilon}\, dy.$$

Since $z(\theta_2^+) < 1+\theta_1^-$, the integral on the right hand side is equal to $G_0(\theta_2^+) + \sqrt{\varepsilon} G_1(\theta_2^+) + o(\sqrt{\varepsilon})$. As $S(1+\theta_1^-) = S(\theta_1^-) - B$ and $-\widehat{V}(\theta_1^+) = -\widehat{V}(\theta_1^-) = S(z(\theta_1^-)) - S(\theta_1^-)$, to complete the proof of the proposition, it remains to show that $S(z(\theta_1^-)) < S(z(\theta_2^+)) + B$. This is clear because $S(z(\theta_1^-)) - B = S(1 + z(\theta_1^-)) = S(z(1+\theta_1^-)) < S(z(\theta_2^+))$. The last inequality follows from the fact that $z(\theta_2^+) < 1+\theta_1^-$. $\square$

Proof of Proposition 6.3. Assume that $z(\theta_1^+) = z(\theta_2^+)$. We estimate the integrals appearing in Assertion 6.A. Clearly,

$$\int_{\theta_1^+}^{\theta_2^-} e^{[S(y)-S(\theta_1^+)]/\varepsilon}\, dy = \int_{\theta_1^+}^{1+\theta_1^+} e^{[S(y)-S(\theta_1^+)]/\varepsilon}\, dy - \int_{\theta_2^-}^{1+\theta_1^+} e^{[S(y)-S(\theta_1^+)]/\varepsilon}\, dy.$$

Since $z(\theta_2^-) = z(\theta_1^+) < 1+\theta_1^+$, in the second integral we may replace $1+\theta_1^+$ by $1+\theta_2^-$. By Proposition 5.2, the right hand side is equal to

$$\begin{aligned}&\big\{G_0\big(\theta_1^+\big) + \sqrt{\varepsilon}G_1\big(\theta_1^+\big) + o(\sqrt{\varepsilon})\big\}e^{[S(z(\theta_1^+))-S(\theta_1^+)]/\varepsilon}\\ &\quad - \big\{G_0\big(\theta_2^-\big) + \sqrt{\varepsilon}G_1\big(\theta_2^-\big) + o(\sqrt{\varepsilon})\big\}e^{[S(z(\theta_2^-))-S(\theta_1^+)]/\varepsilon}.\end{aligned}$$

As $z(\theta_2^-) = z(\theta_1^+)$ and $G_a(\theta_2^-) = G_a(\theta_2^+)$, in view of the definition of $\Delta_{1,2}G_a$, this difference is equal to

$$\big\{(\Delta_{1,2}G_0) + \sqrt{\varepsilon}(\Delta_{1,2}G_1) + o(\sqrt{\varepsilon})\big\}e^{-\widehat{V}(\theta_1^+)/\varepsilon}.$$

On the other hand, since $z(\theta_2^+) = z(\theta_1^-) < 1+\theta_1^-$, there exists a constant $C_0 > 0$, independent of ε, such that

$$\int_{\theta_2^+}^{1+\theta_1^-} e^{[S(y)-S(1+\theta_1^-)]/\varepsilon}\, dy \ge C_0\sqrt{\varepsilon}e^{[S(z(\theta_2^+))-S(1+\theta_1^-)]/\varepsilon}.$$

As $z(\theta_2^+) = z(\theta_1^-)$ and $S(1+\theta_1^-) = S(\theta_1^-) - B$, the expression in the exponential in the previous formula is equal to $-\widehat{V}(\theta_1^-) + B = -\widehat{V}(\theta_1^+) + B$. This proves the first assertion of the proposition.

We turn to the case $z(\theta_2^+) = z(1+\theta_1^+)$. We estimate the integrals appearing in Assertion 6.A. Since $z(\theta_1^+) < z(1+\theta_1^+)$, we have that $z(\theta_1^+) < z(\theta_2^+)$. This implies that $z(\theta_1^+) < \theta_2^-$, so that

$$\int_{\theta_1^+}^{\theta_2^-} e^{[S(y)-S(\theta_1^+)]/\varepsilon}\, dy = \big\{G_0\big(\theta_1^+\big) + G_1\big(\theta_1^+\big)\sqrt{\varepsilon} + o(\sqrt{\varepsilon})\big\}e^{-\widehat{V}(\theta_1^+)/\varepsilon}.$$

On the other hand,

$$\begin{aligned}&\int_{\theta_2^+}^{1+\theta_1^-} e^{[S(y)-S(1+\theta_1^+)]/\varepsilon}\, dy\\ &\quad = \int_{\theta_2^+}^{1+\theta_2^+} e^{[S(y)-S(1+\theta_1^+)]/\varepsilon}\, dy - \int_{1+\theta_1^-}^{1+\theta_2^+} e^{[S(y)-S(1+\theta_1^+)]/\varepsilon}\, dy.\end{aligned}$$

Since $z(1+\theta_1^-) = z(\theta_2^+) < 1+\theta_2^+$ and since $G_a(1+\theta_1^-) = G_a(\theta_1^-)$, $a = 1, 2$, this expression is equal to

$$\begin{aligned}&\big\{G_0\big(\theta_2^+\big) + \sqrt{\varepsilon}G_1\big(\theta_2^+\big) + o(\sqrt{\varepsilon})\big\}e^{[S(z(\theta_2^+))-S(1+\theta_1^+)]/\varepsilon}\\ &\quad - \big\{G_0\big(\theta_1^-\big) + \sqrt{\varepsilon}G_1\big(\theta_1^-\big) + o(\sqrt{\varepsilon})\big\}e^{[S(z(1+\theta_1^-))-S(1+\theta_1^+)]/\varepsilon}.\end{aligned}$$

As $z(\theta_2^+) = z(1+\theta_1^-) = z(1+\theta_1^+)$, and $\widehat{V}(1+\theta_1^+) = \widehat{V}(\theta_1^+)$, the expressions in the exponential above are equal to $-\widehat{V}(\theta_1^+)$, which completes the proof of the proposition because $G_a(\theta_1^-) = G_a(\theta_1^+)$, $a = 0, 1$. □

We conclude this section providing an alternative formula for the capacity.

Assertion 6.B. *Fix two disjoint closed intervals* $\mathcal{A}_1 = [\theta_1^-, \theta_1^+]$, $\mathcal{A}_2 = [\theta_2^-, \theta_2^+]$ *of* $\mathbb{T}$. *Then*,

$$\operatorname{cap}_\varepsilon(\mathcal{A}_1, \mathcal{A}_2) = \varepsilon\big\{h'_{\mathcal{A}_1,\mathcal{A}_2}\big(\theta_1^-\big)m_\varepsilon\big(\theta_1^-\big) - h'_{\mathcal{A}_1,\mathcal{A}_2}\big(\theta_1^+\big)m_\varepsilon\big(\theta_1^+\big)\big\}.$$

Proof. By the expression (6.2) for the Dirichlet form, and since the harmonic function is constant on the sets $\mathcal{A}_1$, $\mathcal{A}_2$, $D_\varepsilon(h_{\mathcal{A}_1,\mathcal{A}_2})$ is equal to the sum of two integrals. The first one is carried over the interval $[\theta_1^+,\theta_2^-]$, while the second one over the interval $[\theta_2^+,\theta_1^-]$. We estimate the first integral. By an integration by parts,

$$\varepsilon\int_{\theta_1^+}^{\theta_2^-}\big[h'_{\mathcal{A}_1,\mathcal{A}_2}(x)\big]^2 m_\varepsilon(x)\,dx = \varepsilon h_{\mathcal{A}_1,\mathcal{A}_2}(x)h'_{\mathcal{A}_1,\mathcal{A}_2}(x)m_\varepsilon(x)\big|_{\theta_1^+}^{\theta_2^-} - \varepsilon\int_{\theta_1^+}^{\theta_2^-} h_{\mathcal{A}_1,\mathcal{A}_2}(x)\partial_x\big\{h'_{\mathcal{A}_1,\mathcal{A}_2}(x)m_\varepsilon(x)\big\}\,dx.$$

Since the harmonic function vanishes at θ_2^- and is equal to 1 at θ_1^+, by (6.1) and since $L_\varepsilon h_{\mathcal{A}_1,\mathcal{A}_2}=0$ on (θ_1^+,θ_2^-), the previous expression is equal to

$$-\varepsilon h'_{\mathcal{A}_1,\mathcal{A}_2}\big(\theta_1^+\big)m_\varepsilon\big(\theta_1^+\big) + \varepsilon R_\varepsilon\int_{\theta_1^+}^{\theta_2^-} h_{\mathcal{A}_1,\mathcal{A}_2}(x)h'_{\mathcal{A}_1,\mathcal{A}_2}(x)\,dx.$$

The integral is equal to $(1/2)\varepsilon R_\varepsilon\{h_{\mathcal{A}_1,\mathcal{A}_2}(\theta_2^-)^2 - h_{\mathcal{A}_1,\mathcal{A}_2}(\theta_1^+)^2\} = -(1/2)\varepsilon R_\varepsilon$.

For similar reasons, the contribution to $D_\varepsilon(h_{\mathcal{A}_1,\mathcal{A}_2})$ of the integral carried over the interval $[\theta_2^+,\theta_1^-]$ is equal to $\varepsilon h'_{\mathcal{A}_1,\mathcal{A}_2}(\theta_1^-)m_\varepsilon(\theta_1^-) + (1/2)\varepsilon R_\varepsilon$. This completes the proof of the assertion. □

7. Metastability among the deepest valleys

We examine in this section the metastable behavior of $X_\varepsilon(t)$ among the deepest valleys. The goal is to define a finite-state, continuous-time Markov chain, called the reduced chain, which describes the evolution of the diffusion $X_\varepsilon(t)$ among the deepest wells in an appropriate time scale.

A similar analysis could be carried out for shallower valleys. This task is left to the interested reader. We assume throughout this section that the drift b satisfies the conditions (H4) of Section 5.4.

7.1. *The deepest valleys*

Denote by $\mathcal{W}_j = (\mathfrak{w}_j^-,\mathfrak{w}_j^+)$, $1\le j\le \boldsymbol{n}$, all valleys of depth H. These are the valleys $\mathcal{V}_i$ introduced in (4.4) such that $\min_{\theta\in\mathcal{V}_i} V(\theta)=0$. Denote by $\mathfrak{m}_{j,k}\in\mathcal{W}_j$, $1\le k\le\kappa(j)$, the global minima of V on $\mathcal{W}_j$. Let $\mathcal{E}_j=(e_j^-,e_j^+)$ be a subset of $\mathcal{W}_j$ which contains all minima $\mathfrak{m}_{j,k}$ and such that $V(e_j^-)=V(e_j^+)<H$, $V(\theta)<V(e_j^-)$ for all $\theta\in\mathcal{E}_j$. The sets $\mathcal{E}_j$ are called wells. We refer to Figure 5 for an illustration. Assume, without loss of generality, that the valleys are ordered in the sense that $0<\mathfrak{m}_{1,1}<\cdots<\mathfrak{m}_{\boldsymbol{n},1}<1$. Denote by $\boldsymbol{\pi}(j)$ the weight of the well $\mathcal{E}_j$:

$$\boldsymbol{\pi}(j)=\sum_{k=1}^{\kappa(j)}\sigma(\mathfrak{m}_{j,k})=\sum_{k=1}^{\kappa(j)}\sqrt{\frac{2\pi}{S''(\mathfrak{m}_{j,k})}},\quad 1\le j\le\boldsymbol{n}, \tag{7.1}$$

where $\sigma(\mathfrak{m}_{j,k})$ has been introduced in (5.6).

Let $\mathfrak{M}_{j,k}$, $1\le k\le\upsilon(j)$, be the global maxima of V which belong to the interval (e_j^+,e_{j+1}^-) and to the landscape which contains $\mathcal{W}_j$. Hence if the valley $\mathcal{W}_j$ is contained in the landscape $[\ell_n,\mathfrak{L}_{n+1}^+]$, $\{\mathfrak{M}_{j,k}:1\le k\le\upsilon(j)\}=\{\mathfrak{M}_l^+ : \mathfrak{M}_l^+\in(e_j^+,e_{j+1}^-)\cap[\ell_n,\mathfrak{L}_{n+1}^+], V(\mathfrak{M}_l^+)=H\}$. We refer to Figure 5 for an illustration. Denote by $\boldsymbol{\sigma}_{j,j+1}$ the sum of the weights of these local maxima:

$$\boldsymbol{\sigma}_{j,j+1}=\sum_{k=1}^{\upsilon(j)}\omega(\mathfrak{M}_{j,k})=\sum_{k=1}^{\upsilon(j)}\sqrt{\frac{2\pi}{-S''(\mathfrak{M}_{j,k})}},\quad 1\le j\le\boldsymbol{n}. \tag{7.2}$$

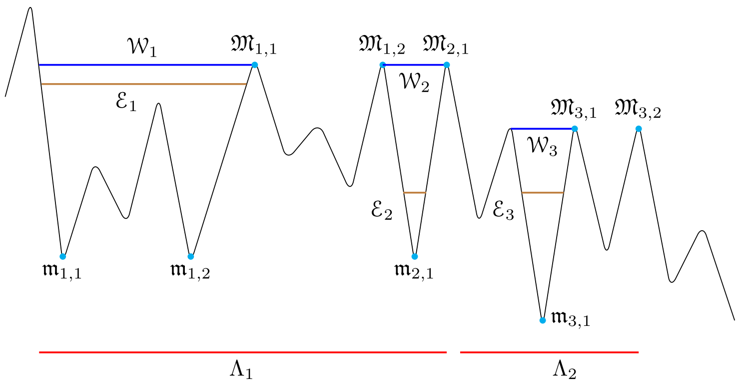

Fig. 5. This figure represents the wells and valleys in two landscapes. There are 3 valleys whose depth is maximal, $\mathcal{W}_1$, $\mathcal{W}_2$ and $\mathcal{W}_3$. The first one contains two global minima of the quasi-potential V, while the other two only one. If the process starts from the well $\mathcal{W}_2$ the next well visited may be either $\mathcal{W}_1$ or $\mathcal{W}_3$, while if it starts from $\mathcal{W}_1$ the next well visited can only be $\mathcal{W}_2$.

7.2. *The evolution among the wells $\mathcal{E}_j$*

The asymptotic behavior of the diffusion $X_\varepsilon(t)$ among the wells $\mathcal{E}_l$ can be foretold. Rename the valleys $\mathcal{W}_1, \dots, \mathcal{W}_{\boldsymbol{n}}$ as $\mathcal{W}_{a,k}$, $1 \le a \le \boldsymbol{p}$, $1 \le k \le n_a$, in such a way that two valleys $\mathcal{W}_{a,k}$, $\mathcal{W}_{a',k'}$ belong to the same landscape if and only if $a = a'$. Denote the minimum $\mathfrak{m}_{l,r}$ by $\mathfrak{m}_{a,k,r}$ if $\mathcal{W}_l = \mathcal{W}_{a,k}$ and assume that the valleys are ordered in the sense that $\mathfrak{m}_{a,k,r} < \mathfrak{m}_{a',k',s}$ if $a < a'$ or if $a = a'$ and $k < k'$.

Assume that $X_\varepsilon(0)$ belongs to $\mathcal{W}_i$. The next visited valley can only be $\mathcal{W}_{i-1}$ or $\mathcal{W}_{i+1}$, where we adopt the convention that $\mathcal{W}_{r\boldsymbol{n}+k} = r + \mathcal{W}_k$. However, if $\mathcal{W}_i = \mathcal{W}_{a,1}$ for some a, since, by Remark 4.4, the diffusion does not visit a landscape to its left, modulo a probability exponentially close to 1, the next visited valley is necessarily $\mathcal{W}_{i+1}$. Hence, if $p(i,j)$ represents the jump probabilities of the reduced chain, we must have that

$$p(i,i+1) + p(i,i-1) = 1 \quad \text{and} \quad p(i,i+1) = 1 \quad \text{if } \mathcal{W}_i = \mathcal{W}_{a,1} \text{ for some } a. \tag{7.3}$$

We may compute the jump probabilities using formula (6.4) for the equilibrium potential. Assume that $\mathcal{W}_i = \mathcal{W}_{a,k}$ for some $k \ge 2$, and let

$$p_\varepsilon(i,i+1) = h_{\mathcal{W}_{i+1},\mathcal{W}_{i-1}}(\mathfrak{m}_{i,1}),$$

where $h_{\mathcal{W}_{i+1},\mathcal{W}_{i-1}}$ is the equilibrium potential introduced in (6.4). An elementary computation gives that

$$p_\varepsilon(i,i+1) = \big[1 + o(1)\big] \frac{\boldsymbol{\sigma}_{i-1,i}}{\boldsymbol{\sigma}_{i-1,i} + \boldsymbol{\sigma}_{i,i+1}}.$$

Therefore, we have to set

$$p(i,i+1) = \frac{\boldsymbol{\sigma}_{i-1,i}}{\boldsymbol{\sigma}_{i-1,i} + \boldsymbol{\sigma}_{i,i+1}}, \quad \text{if } \mathcal{W}_i = \mathcal{W}_{a,k} \text{ for some } k \ge 2. \tag{7.4}$$

Equations (7.3) and (7.4) characterize the jump probabilities of the reduced chain. We turn to the holding rates of the reduced chain. By Proposition 5.3 and Lemma 5.5,

$$\mu_\varepsilon(\mathcal{E}_i) = \big[1 + o(1)\big]\boldsymbol{\mu}(i) \quad \text{where } \boldsymbol{\mu}(i) = \frac{1}{Z} G_1(\mathfrak{m}_{i,1})\boldsymbol{\pi}(i). \tag{7.5}$$

On the other hand, by similar computations to the ones presented in the previous section and by Proposition 5.7, $\mathrm{cap}_\varepsilon(\mathcal{W}_i, \mathcal{W}_{i-1} \cup \mathcal{W}_{i+1}) = [1 + o(1)]e^{-H/\varepsilon}c(i)$, where,

$$c(i) = \frac{G_1(\mathfrak{m}_{i,1})}{Z}\frac{1}{\boldsymbol{\sigma}_{i,i+1}} \quad \text{if } p(i, i+1) = 1,$$
$$c(i) = \frac{G_1(\mathfrak{m}_{i,1})}{Z}\left(\frac{1}{\boldsymbol{\sigma}_{i-1,i}} + \frac{1}{\boldsymbol{\sigma}_{i,i+1}}\right) \quad \text{if } p(i, i+1) < 1.$$

It follows from the previous estimates and from equation (A.8) in [2], that on the time-scale $e^{H/\varepsilon}$, the diffusion $X_\varepsilon(t)$ is expected to evolve among the valleys $\mathcal{W}_j$ as the $\{1, \dots, \boldsymbol{n}\}$-valued, continuous-time Markov chain with jump probabilities given by (7.4) and holding times $\lambda(i)$ given by $\lambda(i) = c(i)/\boldsymbol{\mu}(i)$.

7.3. *The reduced chain*

At this point we have all elements to define the Markov chain which describes the metastable behavior of $X_\varepsilon(t)$. Let $S = \{1, \dots, \boldsymbol{n}\}$ and denote by $R(j, k)$ the jump rates of the continuous-time S-valued Markov chain whose holding rates are λ and whose jump probabilities are p: $R(j, k) = \lambda(j)p(j, k)$, $j \neq k \in S$. By the previous computations,

$$R(j, j+1) = \frac{1}{\boldsymbol{\pi}(j)\boldsymbol{\sigma}_{j,j+1}}, \qquad R(j+1, j) = 0 \quad \text{or} \quad \frac{1}{\boldsymbol{\pi}(j+1)\boldsymbol{\sigma}_{j,j+1}}, \tag{7.6}$$

and $R(j, k) = 0$ if $k \neq j \pm 1$. More precisely, $R(j+1, j) = 0$ if $\mathcal{W}_{j+1} = \mathcal{W}_{a,1}$ for some a, and $R(j+1, j) = [\boldsymbol{\pi}(j+1)\boldsymbol{\sigma}_{j,j+1}]^{-1}$ otherwise.

In view of the previous computation, denote by $\boldsymbol{X}(t)$ the continuous-time Markov chain on S whose generator $\boldsymbol{L}$ is given by

$$(\boldsymbol{L}f)(j) = \sum_{a=\pm 1} R(j, j+a)\big[F(j+a) - F(j)\big]. \tag{7.7}$$

Summation is performed modulo $\boldsymbol{n}$ in the previous formula. The next result is proved at the end of this section.

Lemma 7.1. *The measure $\boldsymbol{\mu}$, introduced in (7.5), is the stationary state of the Markov chain $\boldsymbol{X}(t)$.*

Denote by $D(\mathbb{R}_+, S)$ the space of right-continuous functions $x : \mathbb{R}_+ \to S$ with left-limits endowed with the Skorohod topology, and by $\boldsymbol{Q}_j$, $1 \leq j \leq \boldsymbol{n}$, the probability measure on $D(\mathbb{R}_+, S)$ induced by the Markov process whose generator is $\boldsymbol{L}$ and which starts from j.

7.4. *The metastable behavior*

Denote by $\widehat{X}_\varepsilon(t)$ the process $X_\varepsilon(t)$ speeded-up by $e^{H/\varepsilon}$. This is the diffusion on $\mathbb{T}$ whose generator, denoted by $\widehat{L}_\varepsilon$, is given by $\widehat{L}_\varepsilon = e^{H/\varepsilon}L_\varepsilon$. Let $C(\mathbb{R}_+, \mathbb{T})$ be the space of continuous trajectories $X : \mathbb{R}_+ \to \mathbb{T}$ endowed with the topology of uniform convergence on compact subsets of $\mathbb{R}_+$. Denote by $\mathbb{P}^\varepsilon_\theta$, $\theta \in \mathbb{T}$, the probability measure on $C(\mathbb{R}_+, \mathbb{T})$ induced by the diffusion $\widehat{X}_\varepsilon(t)$ starting from θ. Expectation with respect to $\mathbb{P}^\varepsilon_\theta$ is represented by $\mathbb{E}^\varepsilon_\theta$.

Let

$$\mathcal{E} = \bigcup_{j=1}^{\boldsymbol{n}} \mathcal{E}_j, \qquad \Delta = \mathbb{T} \setminus \mathcal{E}, \qquad \breve{\mathcal{E}}_j = \bigcup_{k:k\neq j} \mathcal{E}_k. \tag{7.8}$$

Denote by $T_\mathcal{E}(t)$, $t \geq 0$, the total time spent by the diffusion $\widehat{X}_\varepsilon(t)$ on the set $\mathcal{E}$ in the time interval $[0, t]$:

$$T_\mathcal{E}(t) := \int_0^t \chi_\mathcal{E}\big(\widehat{X}_\varepsilon(s)\big)\, ds.$$

Denote by $\{S_{\mathscr{E}}(t) : t \geq 0\}$ the generalized inverse of $T_{\mathscr{E}}(t)$:

$$S_{\mathscr{E}}(t) := \sup\big\{s \geq 0 : T_{\mathscr{E}}(s) \leq t\big\}.$$

Clearly, for all $r \geq 0$, $t \geq 0$,

$$\big\{S_{\mathscr{E}}(r) \geq t\big\} = \big\{T_{\mathscr{E}}(t) \leq r\big\}. \tag{7.9}$$

It is also clear that for any starting point $\theta \in \mathbb{T}$, $\lim_{t\to\infty} T_{\mathscr{E}}(t) = \infty$ almost surely. Therefore, the random path $\{Y_\varepsilon(t) : t \geq 0\}$, given by $Y_\varepsilon(t) = \widehat{X}_\varepsilon(S_{\mathscr{E}}(t))$, is well defined for all $t \geq 0$ and takes value in the set $\mathscr{E}$. We call the process $\{Y_\varepsilon(t) : t \geq 0\}$ the trace of $\{\widehat{X}_\varepsilon(t) : t \geq 0\}$ on the set $\mathscr{E}$.

Denote by $\{\mathscr{F}_t^0 : t \geq 0\}$ the natural filtration of $C(\mathbb{R}_+, \mathbb{T})$: $\mathscr{F}_t^0 = \sigma(\widehat{X}_s : 0 \leq s \leq t)$. Fix $\theta_0 \in \mathscr{E}$ and denote by $\{\mathscr{F}_t : t \geq 0\}$ the usual augmentation of $\{\mathscr{F}_t^0 : t \geq 0\}$ with respect to $\mathbb{P}^\varepsilon_{\theta_0}$. We refer to Section III.9 of [16] for a precise definition.

Lemma 7.2. *For each $t \geq 0$, $S_{\mathscr{E}}(t)$ is a stopping time with respect to the filtration $\{\mathscr{F}_t\}$. Let $\{\mathscr{G}_r : r \geq 0\}$ be the filtration given by $\mathscr{G}_r = \mathscr{F}_{S_{\mathscr{E}}(r)}$, and let τ be a stopping time with respect to $\{\mathscr{G}_r\}$. Then, $S_{\mathscr{E}}(\tau)$ is a stopping time with respect to $\{\mathscr{F}_t\}$.*

Proof. Fix $t \geq 0$ and $r \geq 0$. By (7.9),

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

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

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

$$\{\tau \leq t\} \cap \big\{S_{\mathscr{E}}(t) \leq r\big\} \in \mathscr{F}_r.$$

We claim that $\{S_{\mathscr{E}}(\tau) < t\} \in \mathscr{F}_t$. Indeed, by (7.9), this event is equal to $\{T_{\mathscr{E}}(t) > \tau\}$, which can be written as

$$\begin{aligned}\bigcup_q \{\tau \leq q\} \cap \big\{T_{\mathscr{E}}(t) > q\big\} &= \bigcup_q \{\tau \leq q\} \cap \big\{S_{\mathscr{E}}(q) < t\big\} \\ &= \bigcup_q \bigcup_{n\geq 1} \{\tau \leq q\} \cap \big\{S_{\mathscr{E}}(q) \leq t - (1/n)\big\},\end{aligned}$$

where the union is carried over all $q \in \mathbb{Q}$. By the penultimate displayed equation, each term belongs to $\mathscr{F}_{t-(1/n)} \subset \mathscr{F}_t$, which proves the claim.

We may conclude. Since

$$\big\{S_{\mathscr{E}}(\tau) \leq t\big\} = \bigcap_q \big\{S_{\mathscr{E}}(\tau) < t + q\big\},$$

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

Since $S_{\mathscr{E}}(t)$ is a stopping time with respect to the filtration $\{\mathscr{F}_t\}$,

$$Y_\varepsilon(t) = \widehat{X}_\varepsilon\big(S_{\mathscr{E}}(t)\big)$$

is an $\mathcal{E}$-valued, Markov process with respect to the filtration $\mathcal{G}_t = \mathcal{F}_{S_{\mathcal{E}}(t)}$. Let $\Psi : \mathcal{E} \to S = \{1, \dots, \boldsymbol{n}\}$ be the projection given by

$$\Psi(\theta) = \sum_{j=1}^{\boldsymbol{n}} j \chi_{\mathcal{E}_j}(\theta),$$

and denote by $\boldsymbol{x}_\varepsilon(t)$ the projected process which takes value in $S = \{1, \dots, \boldsymbol{n}\}$ and is defined by

$$\boldsymbol{x}_\varepsilon(t) = \Psi\big(Y_\varepsilon(t)\big).$$

Denote by $\mathbb{Q}^\varepsilon_\theta$, $\theta \in \mathcal{E}$, the probability measure on $D(\mathbb{R}_+, \mathcal{E})$ induced by the process $Y_\varepsilon(t)$ starting from θ, and by $\mathbf{Q}^\varepsilon_\theta$ the probability measure on $D(\mathbb{R}_+, S)$ induced by the function Ψ: $\mathbf{Q}^\varepsilon_\theta = \mathbb{Q}^\varepsilon_\theta \circ \Psi^{-1}$. Note that $\mathbf{Q}^\varepsilon_\theta$ corresponds to the distribution of $\boldsymbol{x}_\varepsilon(t)$ starting from $\Psi(\theta)$.

Theorem 7.3. *Fix $1 \le j \le \boldsymbol{n}$ and $\theta_0 \in \mathcal{E}_j$. The sequence of measures $\mathbf{Q}^\varepsilon_{\theta_0}$ converges, as $\varepsilon \to 0$, to the probability measure $\mathbf{Q}_j$ introduced below* (7.7).

Remark 7.4. This is the first example, to our knowledge, that the reduced chain is an non-reversible, irreducible dynamics.

7.5. *Tightness*

The proof of Theorem 7.3 is divided in two steps. We prove in this subsection that the sequence $\mathbf{Q}^\varepsilon_{\theta_0}$ is tight and that all its limit points fulfill certain conditions. In the next subsection, we prove uniqueness of limit points.

Lemma 7.5. *For every $1 \le j \le \boldsymbol{n}$, $\theta_0 \in \mathcal{E}_j$, the sequence of measures $\mathbf{Q}^\varepsilon_{\theta_0}$ is tight. Moreover, every limit point $\mathbf{Q}^*$ of the sequence $\mathbf{Q}^\varepsilon_{\theta_0}$ is such that*

$$\mathbf{Q}^*\big\{x : x(0) = j\big\} = 1 \quad \textit{and} \quad \mathbf{Q}^*\big\{x : x(t) \ne x(t-)\big\} = 0$$

for every $t > 0$.

Proof. Fix $\theta_0 \in \mathcal{E}$. According to Aldous' criterion, we have to show that for every $\delta > 0$, $R > 0$,

$$\lim_{a_0 \to 0} \limsup_{\varepsilon \to 0} \sup \mathbb{Q}^\varepsilon_{\theta_0}\big[\big|\Psi\big(Y(\tau + a)\big) - \Psi\big(Y(\tau)\big)\big| > \delta\big] = 0,$$

where the supremum is carried over all stopping times τ bounded by R and all $0 \le a < a_0$.

By definition of the measure $\mathbb{Q}^\varepsilon_{\theta_0}$ and since $|\Psi(Y_\varepsilon(\tau + a)) - \Psi(Y_\varepsilon(\tau))| > \delta$ entails that $\Psi(Y_\varepsilon(\tau + a)) \ne \Psi(Y_\varepsilon(\tau))$, the probability appearing in the previous displayed equation is bounded by

$$\mathbb{P}^\varepsilon_{\theta_0}\big[\Psi\big(X\big(S_{\mathcal{E}}(\tau + a)\big)\big) \ne \Psi\big(X\big(S_{\mathcal{E}}(\tau)\big)\big)\big].$$

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

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

$$\int_{S_{\mathcal{E}}(\tau)}^{S_{\mathcal{E}}(\tau)+b} \chi_{\mathcal{E}}\big(X(s)\big)\, ds \le a.$$

Equivalently,

$$\int_{S_{\mathcal{E}}(\tau)}^{S_{\mathcal{E}}(\tau)+b} \chi_{\Delta}\big(X(s)\big)\, ds \ge b - a.$$

By Lemma 7.2, $S_{\mathcal{E}}(\tau)$ is a stopping time for the filtration $\{\mathcal{F}_t\}$. Hence, by the strong Markov property and since $X(S_{\mathcal{E}}(t)) \in \mathcal{E}$ for all $t \ge 0$,

$$\mathbb{P}^{\varepsilon}_{\theta_0}\big[S_{\mathcal{E}}(\tau+a) - S_{\mathcal{E}}(\tau) > b\big] \le \sup_{\theta\in\mathcal{E}} \mathbb{P}^{\varepsilon}_{\theta}\bigg[\int_0^b \chi_{\Delta}\big(X(s)\big)\, ds \ge b-a\bigg].$$

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

$$\frac{1}{b-a}\sup_{\theta\in\mathcal{E}} \mathbb{E}^{\varepsilon}_{\theta}\bigg[\int_0^b \chi_{\Delta}\big(X(s)\big)\, ds\bigg] \le \frac{1}{a_0}\sup_{\theta\in\mathcal{E}} \mathbb{E}^{\varepsilon}_{\theta}\bigg[\int_0^{2a_0} \chi_{\Delta}\big(X(s)\big)\, ds\bigg].$$

By Lemma 8.5, this expression vanishes as $\varepsilon \to 0$ for every $a_0 > 0$.

We turn to the case $\{S_{\mathcal{E}}(\tau+a) - S_{\mathcal{E}}(\tau) \le b\}$. On this set we have that $\{\Psi(X(S_{\mathcal{E}}(\tau+a))) \neq \Psi(X(S_{\mathcal{E}}(\tau)))\}$ is contained in $\{\Psi(X(S_{\mathcal{E}}(\tau)+c)) \neq \Psi(X(S_{\mathcal{E}}(\tau)))$ for some $0 \le c \le b\}$. By Lemma 7.2, since $S_{\mathcal{E}}(\tau)$ is a stopping time for the filtration $\{\mathcal{F}_t\}$ and since $X(S_{\mathcal{E}}(t))$ belongs to $\mathcal{E}$ for all t,

$$\begin{aligned} &\mathbb{P}^{\varepsilon}_{\theta_0}\big[\Psi\big(X\big(S_{\mathcal{E}}(\tau+a)\big)\big) \neq \Psi\big(X\big(S_{\mathcal{E}}(\tau)\big)\big), S_{\mathcal{E}}(\tau+a) - S_{\mathcal{E}}(\tau) \le b\big] \\ &\quad\le \sup_{\theta\in\mathcal{E}} \mathbb{P}^{\varepsilon}_{\theta}\big[\Psi\big(X(c)\big) \neq \Psi(\theta) \text{ for some } 0 \le c \le b\big]. \end{aligned}$$

If $\theta \in \mathcal{E}_j$, this later event corresponds to the event $\{H(\breve{\mathcal{E}}_j) \le b\}$, where $\breve{\mathcal{E}}_j = \bigcup_{k\neq j} \mathcal{E}_k$. The supremum is thus bounded by

$$\max_{1\le j\le \boldsymbol{n}} \sup_{\theta\in\mathcal{E}_j} \mathbb{P}^{\varepsilon}_{\theta}\big[H(\breve{\mathcal{E}}_j) \le b\big] = \max_{1\le j\le \boldsymbol{n}} \sup_{\theta\in\mathcal{E}_j} \mathbb{P}^{\varepsilon}_{\theta}\big[H(\breve{\mathcal{E}}_j) \le 2a_0\big].$$

By Corollary 8.4, this expression vanishes as $\varepsilon \to 0$ and then $a_0 \to 0$. This completes the proof of the tightness.

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

$$\lim_{a_0\to 0} \limsup_{\varepsilon\to 0} \mathbf{Q}^{\varepsilon}_{\theta_0}\big[x(t-a) \neq x(t) \text{ for some } 0 \le a \le a_0\big] = 0.$$

Hence, if $\mathbf{Q}^*$ is a limit point of the sequence $\mathbf{Q}^{\varepsilon}_{\theta_0}$,

$$\lim_{a_0\to 0} \mathbf{Q}^*\big[x(t-a) \neq x(t) \text{ for some } 0 \le a \le a_0\big] = 0.$$

This completes the proof of the second assertion of the lemma since $\{x : x(t) \neq x(t-)\} \subset \{x : x(t-a) \neq x(t)$ for some $0 \le a \le a_0\}$ for all $a_0 > 0$. The claim that $\mathbf{Q}^*\{x : x(0) = j\} = 1$ is clear. □

7.6. *Uniqueness of limit points*

The proof of the uniqueness of limit points of the sequence $\mathbf{Q}^{\varepsilon}_{\theta}$ relies on a PDE approach to metastability [6,17].

Lemma 7.6. *Fix $1 \le j \le \boldsymbol{n}$ and $\theta_0 \in \mathcal{E}_j$. Let $\mathbf{Q}^*$ be a limit point of the sequence $\mathbf{Q}^{\varepsilon}_{\theta_0}$. Then, under $\mathbf{Q}^*$, for every $\boldsymbol{F} : S \to \mathbb{R}$,*

$$\boldsymbol{F}\big(x(t)\big) - \int_0^t (\boldsymbol{L}\boldsymbol{F})\big(x(s)\big)\, ds$$

is a martingale.

Proof. Fix $1 \le j \le \boldsymbol{n}$, $\theta_0 \in \mathcal{E}_j$ and a function $\boldsymbol{F} : S \to \mathbb{R}$. Let $f_\varepsilon : \mathbb{T} \to \mathbb{R}$ be the function given by Proposition 9.1. By this result,

$$M_\varepsilon(t) = f_\varepsilon\big(\widehat{X}_\varepsilon(t)\big) - \int_0^t (\widehat{L} f_\varepsilon)\big(\widehat{X}_\varepsilon(s)\big)\, ds = f_\varepsilon\big(\widehat{X}_\varepsilon(t)\big) - \int_0^t \overline{g}_\varepsilon\big(\widehat{X}_\varepsilon(s)\big)\, ds$$

is a martingale with respect to the filtration $\mathcal{F}_t$ and the measure $\mathbb{P}^{\varepsilon}_{\theta_0}$. Since $\{S_{\mathcal{E}}(t) : t \geq 0\}$ are stopping times with respect to $\mathcal{F}_t$,

$$\widehat{M}_{\varepsilon}(t) = M_{\varepsilon}\big(S_{\mathcal{E}}(t)\big) = f_{\varepsilon}\big(Y_{\varepsilon}(t)\big) - \int_0^{S_{\mathcal{E}}(t)} \overline{g}_{\varepsilon}\big(\widehat{X}_{\varepsilon}(s)\big)\, ds$$

is a martingale with respect to $\mathcal{G}_t$. Since $\overline{g}_{\varepsilon}$ vanishes on $\mathcal{E}^c$, by a change of variables,

$$\int_0^{S_{\mathcal{E}}(t)} \overline{g}_{\varepsilon}\big(\widehat{X}_{\varepsilon}(s)\big)\, ds = \int_0^{S_{\mathcal{E}}(t)} \chi_{\mathcal{E}}\big(\widehat{X}_{\varepsilon}(s)\big)\overline{g}_{\varepsilon}\big(\widehat{X}_{\varepsilon}(s)\big)\, ds = \int_0^{t} \overline{g}_{\varepsilon}\big(\widehat{X}_{\varepsilon}\big(S_{\mathcal{E}}(s)\big)\big)\, ds.$$

Hence,

$$\widehat{M}_{\varepsilon}(t) = f_{\varepsilon}\big(Y_{\varepsilon}(t)\big) - \int_0^{t} \overline{g}_{\varepsilon}\big(Y_{\varepsilon}(s)\big)\, ds$$

is a $\{\mathcal{G}_t\}$-martingale under the measure $\mathbb{Q}^{\varepsilon}_{\theta_0}$.

Since $\lim_{\varepsilon\to 0} r(\varepsilon) = 0$, $\overline{g}_{\varepsilon} - g$ vanishes uniformly in $\mathcal{E}$ as $\varepsilon \to 0$. By Proposition 9.1, the same holds for $\widehat{f_{\varepsilon}} - f$. Hence, since $Y_{\varepsilon}(s) \in \mathcal{E}$ for all $s \geq 0$, we may replace in the previous equation g_{ε}, f_{ε} by g, f, respectively, at a cost which vanishes as $\varepsilon \to 0$. Therefore,

$$\widehat{M}_{\varepsilon}(t) = f\big(Y_{\varepsilon}(t)\big) - \int_0^{t} g\big(Y_{\varepsilon}(s)\big)\, ds + o(1)$$

is a $\{\mathcal{G}_t\}$-martingale under the measure $\mathbb{Q}^{\varepsilon}_{\theta_0}$.

Since f and g are constant on each set $\mathcal{E}_j$ with values given by $\boldsymbol{F}$, $\boldsymbol{G}$, respectively, $f(Y_{\varepsilon}(t)) = \boldsymbol{F}(\boldsymbol{x}_{\varepsilon}(t))$, $g(Y_{\varepsilon}(t)) = \boldsymbol{G}(\boldsymbol{x}_{\varepsilon}(t))$. By Lemma 7.5, all limit points of the sequence $\mathbf{Q}^{\varepsilon}_{\theta_0}$ are concentrated on trajectories which are continuous at any fixed time with probability 1. We may, therefore, pass to the limit and conclude that $\boldsymbol{F}(x(t)) - \int_0^t (\boldsymbol{LF})(x(s))\, ds$ is a martingale under $\mathbf{Q}^*$. □

Proof of Theorem 7.3. The assertion is a consequence of Lemma 7.5, Lemma 7.6 and the fact that there is only one measure $\mathbf{Q}$ on $D(\mathbb{R}_+, S)$ such that $\mathbf{Q}[x(0) = j] = 1$ and

$$\boldsymbol{F}\big(x(t)\big) - \int_0^t (\boldsymbol{LF})\big(x(s)\big)\, ds$$

is a martingale for all $\boldsymbol{F} : S \to \mathbb{R}$. □

7.7. *Proof of Lemma* 7.1

We have seen in Section 7.3 that the jump rates depend on the position of the valley in the landscape. If the valley is the left-most valley, it jumps only to the right. We need therefore a notation to indicate if a point $j \in S$ is the index of a left-most valley or not.

Recall that the wells $\mathcal{W}_j$ which belong to the same landscape are represented as $\mathcal{W}_{a,1}, \dots, \mathcal{W}_{a,n_a}$. We may thus associate each $j \in S$ to a pair (a, ℓ), where $a \in \{1, \dots, \boldsymbol{p}\}$ represents the landscape and $\ell \in \{1, \dots, n_a\}$ the position in the landscape. Hence, S can also be written as

$$S = \big\{(1,1), \dots, (1, n_1), \dots, (\boldsymbol{p}, 1), \dots, (\boldsymbol{p}, n_{\boldsymbol{p}})\big\}.$$

Consider the subset $S_a = \{(a,1), \dots, (a, n_a)\}$ of S. Recall from (7.2) the notation $\boldsymbol{\sigma}[j, j+1] = \boldsymbol{\sigma}_{j,j+1}$. For $1 \leq j < n_a$, the Markov chain $\boldsymbol{X}(t)$ defined in Section 7.3 jumps from (a, j) to $(a, j+1)$ at rate $\{\boldsymbol{\pi}(a, j)\boldsymbol{\sigma}[(a, j), (a, j+1)]\}^{-1}$ and from $(a, j+1)$ to (a, j) at rates $\{\boldsymbol{\pi}(a, j+1)\boldsymbol{\sigma}[(a, j), (a, j+1)]\}^{-1}$. Additionally, it jumps from (a, n_a) to $(a+1, 1)$. If we disregard this last jump, on the set S_a, the Markov chain behaves as a reversible Markov chain whose equilibrium state is $\boldsymbol{\pi}$. The additional jump from (a, n_a) to $(a+1, 1)$ changes the stationary state by the multiplicative factor $G_1(\mathfrak{m}_{a,j,1})$. This is the content of Lemma 7.1.

Proof of Lemma 7.1. Consider a function $F\colon S\to\mathbb{R}$, and recall that a point $j\in S$ is also represented as (a,k). With this notation, $E_{\boldsymbol{\mu}}[\boldsymbol{L}F]$ becomes

$$\sum_{a=1}^{\boldsymbol{p}}\boldsymbol{\mu}(a,n_a)\frac{1}{\boldsymbol{\pi}(a,n_a)\boldsymbol{\sigma}[(a,n_a),(a+1,1)]}\big[F(a+1,1)-F(a,n_a)\big]$$
$$+\sum_{a=1}^{\boldsymbol{p}}\sum_{j=1}^{n_a-1}\boldsymbol{\mu}(a,j)\frac{1}{\boldsymbol{\pi}(a,j)\boldsymbol{\sigma}[(a,j),(a,j+1)]}\big[F(a,j+1)-F(a,j)\big]$$
$$+\sum_{a=1}^{\boldsymbol{p}}\sum_{j=1}^{n_a-1}\boldsymbol{\mu}(a,j+1)\frac{1}{\boldsymbol{\pi}(a,j+1)\boldsymbol{\sigma}[(a,j),(a,j+1)]}\big[F(a,j)-F(a,j+1)\big].$$

Since the first summation is performed modulo $\boldsymbol{p}$ and since, by (7.5), $\boldsymbol{\mu}(a,j)/\boldsymbol{\pi}(a,j)=(1/Z)G_1(\mathfrak{m}_{a,j,1})$, a change of variables yields that the first sum can be rewritten as

$$\frac{1}{Z}\sum_{a=1}^{\boldsymbol{p}}\frac{G_1(\mathfrak{m}_{a-1,n_{a-1},1})}{\boldsymbol{\sigma}[(a-1,n_{a-1}),(a,1)]}F(a,1)-\frac{1}{Z}\sum_{a=1}^{\boldsymbol{p}}\frac{G_1(\mathfrak{m}_{a,n_a,1})}{\boldsymbol{\sigma}[(a,n_a),(a+1,1)]}F(a,n_a).$$

By definition of G_1 and $\boldsymbol{\sigma}_{j,j+1}$, the previous ratios are equal to 1 and the difference becomes

$$\frac{1}{Z}\sum_{a=1}^{\boldsymbol{p}}F(a,1)-\frac{1}{Z}\sum_{a=1}^{\boldsymbol{p}}F(a,n_a). \tag{7.10}$$

Use the identity $\boldsymbol{\mu}(a,j)/\boldsymbol{\pi}(a,j)=(1/Z)G_1(\mathfrak{m}_{a,j,1})$ to rewrite the last two terms of the first displayed formula of this proof as

$$\frac{1}{Z}\sum_{a=1}^{\boldsymbol{p}}\sum_{j=2}^{n_a}F(a,j)\left\{\frac{G_1(\mathfrak{m}_{a,j-1,1})}{\boldsymbol{\sigma}[(a,j-1),(a,j)]}-\frac{G_1(\mathfrak{m}_{a,j,1})}{\boldsymbol{\sigma}[(a,j-1),(a,j)]}\right\}$$
$$+\frac{1}{Z}\sum_{a=1}^{\boldsymbol{p}}\sum_{j=1}^{n_a-1}F(a,j)\left\{\frac{G_1(\mathfrak{m}_{a,j+1,1})}{\boldsymbol{\sigma}[(a,j),(a,j+1)]}-\frac{G_1(\mathfrak{m}_{a,j,1})}{\boldsymbol{\sigma}[(a,j),(a,j+1)]}\right\}.$$

By definition of G_1 and $\boldsymbol{\sigma}$, $G_1(\mathfrak{m}_{a,j-1,1})-G_1(\mathfrak{m}_{a,j,1})=\boldsymbol{\sigma}[(a,j-1),(a,j)]$. This sum is thus equal to

$$\frac{1}{Z}\sum_{a=1}^{\boldsymbol{p}}\sum_{j=2}^{n_a}F(a,j)-\frac{1}{Z}\sum_{a=1}^{\boldsymbol{p}}\sum_{j=1}^{n_a-1}F(a,j)=\frac{1}{Z}\sum_{a=1}^{\boldsymbol{p}}\big\{F(a,n_a)-F(a,1)\big\}.$$

This term cancels (7.10), which completes the proof of the assertion. □

The same proof yields the next claim, which is needed later.

Lemma 7.7. *Fix a function $F:S\to\mathbb{R}$. For every $1\le a\le p$ and every $1\le \ell\le n_a$,*

$$F(a,\ell)-F(1,1)=\sum_{b=1}^{a-1}\sum_{k=1}^{n_b}(\boldsymbol{L}F)(b,k)\boldsymbol{\pi}(b,k)G_1(\mathfrak{m}_{b,k,1})$$
$$+\sum_{k=1}^{\ell-1}(\boldsymbol{L}F)(a,k)\boldsymbol{\pi}(a,k)\big[G_1(\mathfrak{m}_{a,k,1})-G_1(\mathfrak{m}_{a,\ell,1})\big].$$

8. Hitting times estimates via enlarged processes

We prove in this section an upper bound for the probability of the transition time between wells to be small. This estimate plays a central role in the proof of the tightness of a sequence of metastable processes. The argument presented below is absolutely general and does not rely on the one-dimensionality of the process.

The argument is based on an enlargement of the process $X_\varepsilon(t)$. Fix $\gamma > 0$, let $\mathbb{T}_2 = \mathbb{T} \times \{-1, 1\}$, and consider the process $\boldsymbol{X}^\gamma_\varepsilon(t) = (X^\gamma_\varepsilon(t), \sigma(t))$ on $\mathbb{T}_2$ whose generator L^γ_ε is given by

$$\big(L^\gamma_\varepsilon f\big)(\theta,\sigma) = (\widehat{L}_\varepsilon f)(\theta,\sigma) + \gamma\big[f(\theta,-\sigma) - f(\theta,\sigma)\big].$$

In the first term on the right hand side, the derivatives act only on the first coordinate. The process $\boldsymbol{X}^\gamma_\varepsilon(t)$ is named the enlarged process. The first coordinate evolves as the original process, while the second one, independently from the first, jumps from ± 1 to ∓ 1 at rate γ.

Denote by $\mathbb{P}^{\gamma,\varepsilon}_{(\theta,\sigma)}$ the probability measure on $D(\mathbb{R}_+, \mathbb{T}_2)$ induced by the Markov process $\boldsymbol{X}^\gamma_\varepsilon$ starting from (θ,σ). It is clear that the measure μ^γ_ε, given by

$$\int_{\mathbb{T}_2} f(\theta,\sigma)\mu^\gamma_\varepsilon(d\theta,d\sigma) = \frac{1}{2}\int_{\mathbb{T}} f(\theta,1)\mu_\varepsilon(d\theta) + \frac{1}{2}\int_{\mathbb{T}} f(\theta,-1)\mu_\varepsilon(d\theta),$$

is the unique stationary state of the process $\boldsymbol{X}^\gamma_\varepsilon$.

Fix an open interval I of $\mathbb{T}$, and let $(I^c,1) = \{(\theta,\sigma) \in \mathbb{T}_2 : \theta \in I^c, \sigma = 1\}$, $(\mathbb{T},-1) = \{(\theta,\sigma) \in \mathbb{T}_2 : \sigma = -1\}$. Denote by $\widehat{h}_I : \mathbb{T}_2 \to \mathbb{R}_+$ the equilibrium potential between $(I^c,1)$ and $(\mathbb{T},-1)$:

$$\widehat{h}_I(\theta,\sigma) = \mathbb{P}^{\gamma,\varepsilon}_{(\theta,\sigma)}[H_{(I^c,1)} \le H_{(\mathbb{T},-1)}]. \tag{8.1}$$

Clearly, $\widehat{h}_I(\theta,-1) = 0$ for all $\theta \in \mathbb{T}$. Let $h_I : \mathbb{T} \to \mathbb{R}_+$ be given by $h_I(\theta) = \widehat{h}_I(\theta,1)$, and denote by $\mathrm{cap}_{\gamma,\varepsilon}[(I^c,1), (\mathbb{T},-1)]$ the capacity between the sets $(I^c,1)$, $(\mathbb{T},-1)$, which is given by the energy of h_I:

$$\mathrm{cap}_{\gamma,\varepsilon}\big[\big(I^c,1\big),(\mathbb{T},-1)\big] = \frac{1}{2}\varepsilon e^{H/\varepsilon}\int_I \big(\partial_\theta h_I(\theta)\big)^2\mu_\varepsilon(d\theta) + \frac{\gamma}{2}\int_I h_I(\theta)^2\mu_\varepsilon(d\theta). \tag{8.2}$$

Proposition 8.1. *Let I be an open interval of $\mathbb{T}$, $\theta \in I$. Then, for every $A > 0$, $\mathfrak{m} \in I$ and $\eta > 0$ such that $(\mathfrak{m}-\eta, \mathfrak{m}+\eta) \subset I$,*

$$\begin{aligned}\mathbb{P}^\varepsilon_\theta[H_{I^c} \le A] \le & \sup_{\mathfrak{m}-\eta \le \theta' \le \mathfrak{m}+\eta} \mathbb{P}^\varepsilon_\theta[H_{I^c} < H_{\theta'}] \\ & + \frac{2eA}{\mu_\varepsilon(\mathfrak{m}-\eta,\mathfrak{m}+\eta)}\mathrm{cap}_{\gamma,\varepsilon}\big[\big(I^c,1\big),(\mathbb{T},-1)\big],\end{aligned}$$

where $\gamma = A^{-1}$.

Remark 8.2. We will select later I as a valley and $\mathfrak{m}$ as a minimum in I.

Remark 8.3. Let I be an open interval of $\mathbb{T}$. The same arguments show that for every $A > 0$, $J \subset I$,

$$\frac{1}{\mu_\varepsilon(J)}\int_J \mathbb{P}^\varepsilon_\theta[H_{I^c} \le A]\mu_\varepsilon(d\theta) \le \frac{2eA}{\mu_\varepsilon(J)}\mathrm{cap}_{\gamma,\varepsilon}\big[\big(I^c,1\big),(\mathbb{T},-1)\big],$$

where $\gamma = A^{-1}$. But the proof does not use the one-dimensionality of the process, that is, the fact that the process visit points. We leave the proof of this remark to the reader.

The proof of Proposition 8.1 relies on an idea taken from [3], and it is divided in several assertions.

Assertion 8.A. *Let I be an open interval of $\mathbb{T}$, $\theta \in I$. Fix $A > 0$, and let $\mathfrak{e}_A$ be a mean-A, exponential random variable independent of the process $\widehat{X}_\varepsilon$. Then,*

$$\mathbb{P}^\varepsilon_\theta[H_{I^c} \le A] \le e\mathbb{P}^\varepsilon_\theta[H_{I^c} \le \mathfrak{e}_A].$$

Proof. By independence, if $\gamma = A^{-1}$,

$$\mathbb{P}^\varepsilon_\theta[H_{I^c} \le \mathfrak{e}_A] \ge \int_A^\infty \mathbb{P}^\varepsilon_\theta[H_{I^c} \le t]\gamma e^{-\gamma t}\, dt \ge \mathbb{P}^\varepsilon_\theta[H_{I^c} \le A]\int_A^\infty \gamma e^{-\gamma t}\, dt,$$

as claimed. □

To estimate $\mathbb{P}^\varepsilon_\theta[H_{I^c} \le \mathfrak{e}_A]$ in terms capacities, we interpret the exponential time $\mathfrak{e}_A$ as the time the process $\boldsymbol{X}^\gamma_\varepsilon(t)$ starting from $(\theta, 1)$ jumps to $(\mathbb{T}, -1)$ provided $\gamma = A^{-1}$. Indeed, since the second coordinate jumps at rate γ, independently from the first one, for any open interval I of $\mathbb{T}$ and any $\theta \in I$,

$$\mathbb{P}^\varepsilon_\theta[H_{I^c} \le \mathfrak{e}_A] = \mathbb{P}^{\gamma,\varepsilon}_{(\theta,1)}[H_{(I^c,1)} \le H_{(\mathbb{T},-1)}] = h_I(\theta). \tag{8.3}$$

Assertion 8.B. *Let I be an open interval of $\mathbb{T}$, $\theta \in I$. Then,*

$$\gamma \int_I h_I(\theta)\mu_\varepsilon(d\theta) = 2\mathrm{cap}_{\gamma,\varepsilon}\big[\big(I^c, 1\big), (\mathbb{T}, -1)\big].$$

Proof. The function $H(\theta, \sigma) = h_I(\theta)\mathbf{1}\{\sigma = 1\}$ is harmonic on $(I, 1)$, so that $L^\gamma_\varepsilon H = 0$ on this set. Multiplying this identity by $1 - h_I$, integrating over $(I, 1)$ with respect to μ^γ_ε, and integrating by parts yields that

$$\begin{aligned} 0 &= \varepsilon e^{H/\varepsilon}\int_I \big(h_I'\big)^2 m_\varepsilon\, d\theta - \varepsilon e^{H/\varepsilon}\int_I (1-h_I)h_I' m_\varepsilon'\, d\theta \\ &\quad + e^{H/\varepsilon}\int_I (1-h_I)h_I' b m_\varepsilon\, d\theta - \gamma\int_I (1-h_I)h_I m_\varepsilon\, d\theta. \end{aligned}$$

By (6.1), $\varepsilon m_\varepsilon' - b m_\varepsilon$ is equal to a constant, denoted below by $-\varepsilon R_\varepsilon$. Hence, if $I = (u, v)$, the sum of the second and third terms of the previous equation is equal to

$$\varepsilon R_\varepsilon e^{H/\varepsilon}\int_I (1-h_I)h_I'\, d\theta = -\frac{1}{2}\varepsilon R_\varepsilon e^{H/\varepsilon}\big\{\big[1-h_I(v)\big]^2 - \big[1-h_I(u)\big]^2\big\} = 0,$$

because $h_I(u) = h_I(v) = 1$. This proves the assertion in view of formula (8.2) for the capacity. □

In the next assertion we take advantage of working in a one-dimensional space. More precisely, although the next statement is correct in higher dimension, it is empty since the first term on the right-hand side is equal to 1.

Assertion 8.C. *Let I be an open interval of $\mathbb{T}$. For every $\theta, \theta' \in I$, $A > 0$,*

$$\mathbb{P}^\varepsilon_\theta[H_{I^c} \le A] \le \mathbb{P}^\varepsilon_\theta[H_{I^c} < H_{\theta'}] + \mathbb{P}^\varepsilon_{\theta'}[H_{I^c} \le A].$$

Proof. Intersect the set $\{H_{I^c} \le A\}$ with the event $\{H_{I^c} < H_{\theta'}\}$ and its complement. The first set appears on the right hand side. On the other hand, on $\{H_{\theta'} < H_{I^c}\}$, $H_{I^c} = H_{I^c} \circ \vartheta(H_{\theta'}) + H_{\theta'}$, where $\vartheta(t)$ represents the translation of a trajectory by t. In particular, $H_{I^c} \le A$ implies that $H_{I^c} \circ \vartheta(H_{\theta'}) \le A$. Hence, by the strong Markov property,

$$\begin{aligned} \mathbb{P}^\varepsilon_\theta[H_{\theta'} < H_{I^c}, H_{I^c} \le A] &\le \mathbb{P}^\varepsilon_\theta\big[H_{\theta'} < H_{I^c}, H_{I^c} \circ \vartheta(H_{\theta'}) \le A\big] \\ &= \mathbb{E}^\varepsilon_\theta\big[\mathbf{1}\{H_{\theta'} < H_{I^c}\}\mathbb{P}^\varepsilon_{\theta'}[H_{I^c} \le A]\big] \le \mathbb{P}^\varepsilon_{\theta'}[H_{I^c} \le A]. \end{aligned}$$

This proves the claim. □

We are now in a position to prove Proposition 8.1.

Proof of Proposition 8.1. By Assertion 8.C,

$$\mathbb{P}^{\varepsilon}_{\theta}[H_{I^c} \le A] \le \sup_{\mathfrak{m}-\eta\le\theta'\le\mathfrak{m}+\eta} \mathbb{P}^{\varepsilon}_{\theta}[H_{I^c} < H_{\theta'}] + \frac{1}{\mu_\varepsilon(\mathfrak{m}-\eta,\mathfrak{m}+\eta)} \int_{\mathfrak{m}-\eta}^{\mathfrak{m}+\eta} \mathbb{P}^{\varepsilon}_{\theta'}[H_{I^c} \le A] m_\varepsilon\big(\theta'\big)\, d\theta'.$$

By Assertions 8.A and 8.B and (8.3), the integral appearing in the second term is less than or equal to

$$e\int_{\mathfrak{m}-\eta}^{\mathfrak{m}+\eta} \mathbb{P}^{\varepsilon}_{\theta'}[H_{I^c} \le \mathfrak{e}_A] m_\varepsilon\big(\theta'\big)\, d\theta' \le e\int_I \mathbb{P}^{\varepsilon}_{\theta'}[H_{I^c} \le \mathfrak{e}_A] m_\varepsilon\big(\theta'\big)\, d\theta' \le \frac{2e}{\gamma}\mathrm{cap}_{\gamma,\varepsilon}\big[\big(I^c,1\big),(\mathbb{T},-1)\big]$$

which completes the proof of the proposition. □

We apply Proposition 8.1 to the case in which the interval I is a valley $\mathcal{W}_j$, $1\le j\le \boldsymbol{n}$, introduced in Section 7.1. Denote by $d(\theta,J)$ the distance from θ to a subset J of $\mathbb{T}$: $d(\theta,J)=\inf\{d(\theta,\theta'):\theta'\in J\}$, where d represents the distance in the torus. Recall from Section 7.1 the definition of a well $\mathcal{E}_j$.

Corollary 8.4. *Fix* $1\le j\le\boldsymbol{n}$. *Then*,

$$\lim_{a\to 0}\limsup_{\varepsilon\to 0}\sup_{\theta\in\mathcal{E}_j}\mathbb{P}^{\varepsilon}_{\theta}[H_{\mathcal{W}_j^c}\le a]=0.$$

Proof. Let $\mathfrak{m}=\mathfrak{m}_{j,1}$ and fix $\eta>0$ such that $(\mathfrak{m}-\eta,\mathfrak{m}+\eta)\subset\mathcal{E}_j$. We need to estimate the two terms which appear on the right hand side of Proposition 8.1. On the one hand, it follows from the explicit formulae for the equilibrium potential derived in Section 6 that

$$\limsup_{\varepsilon\to 0}\sup_{\theta\in\mathcal{E}_j}\sup_{\mathfrak{m}-\eta\le\theta'\le\mathfrak{m}+\eta}\mathbb{P}^{\varepsilon}_{\theta}[H_{\mathcal{W}_j^c}<H_{\theta'}]=0.$$

On the other hand, since $V(\mathfrak{m})=0$, there exists a constant $c(\eta)>0$, independent of ε, such that $\mu_\varepsilon(\mathfrak{m}-\eta,\mathfrak{m}+\eta)\ge c(\eta)$. It remains, therefore, to show that

$$\lim_{a\to 0}\limsup_{\varepsilon\to 0} a\ \mathrm{cap}_{\gamma,\varepsilon}\big[\big(\mathcal{W}_j^c,1\big),(\mathbb{T},-1)\big], \tag{8.4}$$

where $\gamma=a^{-1}$.

Since the process is interrupted as it reaches the boundary of the valley $\mathcal{W}_j$, it evolves as a reversible process, and all computations can be performed with respect to this later one.

On the set of functions $f:I\to\mathbb{R}$ which are equal to 1 at the boundary of I, the energy which appears on the right hand side of (8.2) is minimized by the equilibrium potential h_I introduced in (8.1). Hence, in order to prove (8.4), it is enough to exhibit a function $f_\varepsilon:\mathcal{W}_j\to\mathbb{R}$ which is equal to 1 at the boundary of $\mathcal{W}_j$ and such that

$$\lim_{a\to 0}\limsup_{\varepsilon\to 0}\left\{ae^{H/\varepsilon}\varepsilon\int_{\mathcal{W}_j}\big(\partial_\theta f_\varepsilon(\theta)\big)^2 m_\varepsilon(\theta)\,d\theta+\int_{\mathcal{W}_j} f_\varepsilon(\theta)^2 m_\varepsilon(\theta)\,d\theta\right\}=0.$$

Let $f_\varepsilon:\mathcal{W}_j\to\mathbb{R}$ be the continuous function defined by

$$f_\varepsilon(\theta)=\frac{\int_{\mathfrak{m}}^{\theta}e^{S(y)/\varepsilon}\,dy}{\int_{\mathfrak{m}}^{\mathfrak{w}_j^+}e^{S(y)/\varepsilon}\,dy}\chi_{[\mathfrak{m},\mathfrak{w}_j^+]}(\theta)+\frac{\int_{\theta}^{\mathfrak{m}}e^{S(y)/\varepsilon}\,dy}{\int_{\mathfrak{w}_j^-}^{\mathfrak{m}}e^{S(y)/\varepsilon}\,dy}\chi_{[\mathfrak{w}_j^-,\mathfrak{m}]}(\theta).$$

Note that this function is not differentiable at $\mathfrak{m} = \mathfrak{m}_{j,1}$. It is easy to check that this test function fulfills the condition introduced in the penultimate displayed equation, which completes the proof of the corollary. □

We turn to an estimate for the time spent outside the wells. Recall from Section 7.4 the definition of the speeded-up processes $\widehat{X}_\varepsilon(\cdot)$.

Lemma 8.5. *For every* $1 \le j \le \boldsymbol{n}$, $t > 0$,

$$\lim_{\varepsilon\to 0} \sup_{\theta\in\mathcal{E}_j} \mathbb{E}^\varepsilon_\theta\left[\int_0^t \chi_\Delta\big(\widehat{X}_\varepsilon(s)\big)\, ds\right] = 0.$$

Proof. Fix $\eta > 0$ and let $\mathcal{E}_j^{(\eta)} = \{\theta \in \mathcal{E}_j : d(\theta, \Delta) > \eta\}$. The time integral appearing in the statement of the lemma is bounded above by

$$H_{\mathcal{E}_j^{(\eta)}} + \int_{H_{\mathcal{E}_j^{(\eta)}}}^{t+H_{\mathcal{E}_j^{(\eta)}}} \chi_\Delta\big(\widehat{X}_\varepsilon(s)\big)\, ds.$$

As observed in Section 6, in one-dimension diffusions visit points and one can compute the capacity between singletons and sets. It follows from the proof of Proposition 3.3 in [13] that

$$\mathbb{E}^\varepsilon_\theta\big[H_{\mathcal{E}_j^{(\eta)}}\big(\widehat{X}_\varepsilon(\cdot)\big)\big] = \frac{e^{-H/\varepsilon}}{\mathrm{cap}_\varepsilon(\{\theta\}, \mathcal{E}_j^{(\eta)})} \int_{\mathbb{T}} h^*_{\{\theta\},\mathcal{E}_j^{(\eta)}}(\theta')m_\varepsilon(\theta')\, d\theta',$$

where the factor $e^{-H/\varepsilon}$ appeared because the process $\widehat{X}_\varepsilon(s)$ has been speeded-up by $e^{H/\varepsilon}$. In this formula, $h^*_{\{\theta\},\mathcal{E}_j^{(\eta)}}$ represents the equilibrium potential between $\{\theta\}$ and $\mathcal{E}_j^{(\eta)}$ for the adjoint process. Therefore,

$$\mathbb{E}^\varepsilon_\theta\big[H_{\mathcal{E}_j^{(\eta)}}\big(\widehat{X}_\varepsilon(\cdot)\big)\big] \le \frac{e^{-H/\varepsilon}}{\mathrm{cap}_\varepsilon(\{\theta\}, \mathcal{E}_j^{(\eta)})}.$$

As in Section 6, it is possible to derive an explicit formula for this capacity and to show that this expression vanishes as $\varepsilon \to 0$, uniformly in $\theta \in \mathcal{E}_j$.

By the strong Markov property, it remains to show that for every $1 \le j \le \boldsymbol{n}$, $t > 0$, $\eta > 0$,

$$\lim_{\varepsilon\to 0} \sup_{\theta\in\mathcal{E}_j^{(\eta)}} \mathbb{E}^\varepsilon_\theta\left[\int_0^t \chi_\Delta\big(\widehat{X}_\varepsilon(s)\big)\, ds\right] = 0.$$

Let $\eta' > 0$ be such that $\mathfrak{m}_{j,1} \in \mathcal{E}_j^{(2\eta')}$. It follows from the explicit formulae for the equilibrium potentials computed in Section 6 that

$$\lim_{\varepsilon\to 0} \sup_{\theta\in\mathcal{E}_j^{(\eta)}} \sup_{\theta'\in\mathcal{E}_j^{(\eta')}} \mathbb{P}^\varepsilon_\theta[H_\Delta < H_{\theta'}] = 0.$$

Since the time integral appearing in the statement of the lemma is bounded, it follows from the previous estimate that we may insert inside the expectation the indicator of the set $\{H_{\theta'} < H_\Delta\}$. Since we may start the time integral from H_Δ, on the set $\{H_{\theta'} < H_\Delta\}$

$$\int_0^t \chi_\Delta\big(\widehat{X}_\varepsilon(s)\big)\, ds = \int_{H_{\theta'}}^t \chi_\Delta\big(\widehat{X}_\varepsilon(s)\big)\, ds \le \int_0^t \chi_\Delta\big(\widehat{X}_\varepsilon(s)\big)\, ds \circ \vartheta(H_{\theta'}).$$

Hence, by the strong Markov property,

$$\mathbb{E}^{\varepsilon}_{\theta}\Big[\mathbf{1}\{H_{\theta'} < H_{\Delta}\}\int_0^t \chi_{\Delta}\big(\widehat{X}_{\varepsilon}(s)\big)\,ds\Big] \le \mathbb{E}^{\varepsilon}_{\theta'}\Big[\int_0^t \chi_{\Delta}\big(\widehat{X}_{\varepsilon}(s)\big)\,ds\Big].$$

Note that the starting point changed from θ to θ'.

In view of the previous bounds, to prove the lemma it is enough to show that

$$\lim_{\varepsilon\to 0}\frac{1}{\mu_{\varepsilon}([\mathfrak{m}_{j,1}-\eta',\mathfrak{m}_{j,1}+\eta'])}\int_{\mathfrak{m}_{j,1}-\eta'}^{\mathfrak{m}_{j,1}+\eta'}\mathbb{E}^{\varepsilon}_{\theta'}\Big[\int_0^t \chi_{\Delta}\big(\widehat{X}_{\varepsilon}(s)\big)\,ds\Big]\mu_{\varepsilon}\big(d\theta'\big) = 0.$$

Since $V(\mathfrak{m}_{j,1}) = 0$, there exists a constant $c(\eta') > 0$, independent of ε, such that $\mu_{\varepsilon}([\mathfrak{m}_{j,1}-\eta',\mathfrak{m}_{j,1}+\eta']) \ge c(\eta')$. On the other hand, the integral is bounded by

$$\int_{\mathbb{T}}\mathbb{E}^{\varepsilon}_{\theta'}\Big[\int_0^t \chi_{\Delta}\big(\widehat{X}_{\varepsilon}(s)\big)\,ds\Big]\mu_{\varepsilon}\big(d\theta'\big) = t\,\mu_{\varepsilon}(\Delta),$$

which vanishes as $\varepsilon \to 0$. □

9. The Poisson equation

We examine in this section properties of the solution of the equation $\widehat{L}_{\varepsilon} f = g$ for a function $g : \mathbb{T} \to \mathbb{R}$ which has mean zero with respect to μ_{ε}. We assume in this section the conditions (H4) of Section 5.4. Recall the notation introduced in Section 7, and that we denote by $\mathfrak{w}_i^{\pm}$ the endpoints of the well $\mathcal{W}_i$. Throughout this section, we assume, without loss of generality, that $\mathcal{W}_1$ is the left-most valley of a landscape: $\mathcal{W}_1 = \mathcal{W}_{1,1}$ in the notation introduced in the paragraph below (7.2). Therefore, there exists $\eta > 0$ such that

$$S(x) \ge S(\mathfrak{m}_{1,1}) + \eta \quad \text{for all } -\infty < x \le \mathfrak{w}_1^{-}. \tag{9.1}$$

Fix a function $\boldsymbol{F} : S \to \mathbb{R}$, and let $\boldsymbol{G} = \boldsymbol{L}\boldsymbol{F}$. Denote by $g : \mathbb{T} \to \mathbb{R}$ the function given by

$$g = \sum_{1\le i\le \boldsymbol{n}} \boldsymbol{G}(i)\chi_{\mathcal{E}_i}.$$

Assertion 9.A. *We have that* $\lim_{\varepsilon\to 0} E_{\mu_{\varepsilon}}[g] = 0$.

Proof. By definition of the function g,

$$E_{\mu_{\varepsilon}}[g] = \sum_{i=1}^{\boldsymbol{n}} \boldsymbol{G}(i)\mu_{\varepsilon}(\mathcal{E}_i).$$

Fix $1 \le i \le \boldsymbol{n}$. By definition of μ_{ε}, by Proposition 5.7 and since, by Remark 5.1, G_1 is constant on valleys,

$$\mu_{\varepsilon}(\mathcal{E}_i) = \big[1 + o(1)\big]\frac{1}{Z}\frac{1}{\sqrt{\varepsilon}}G_1(\mathfrak{m}_{i,1})\int \chi_{\mathcal{E}_i}(x)e^{-V(x)/\varepsilon}\,dx.$$

By Remark 4.1, V and S differ by an additive constant on valleys. Hence, since $V(\mathfrak{m}_{i,k}) = 0$, the previous expression is equal to

$$\big[1 + o(1)\big]\frac{1}{Z}G_1(\mathfrak{m}_{i,1})\sum_{k=1}^{\kappa(i)}\sigma(\mathfrak{m}_{i,k}) = \big[1 + o(1)\big]\boldsymbol{\mu}(i),$$

where the last identity follows from the definition of $\boldsymbol{\mu}$ given in (7.5).

In conclusion,

$$E_{\mu_\varepsilon}[g] = \big[1 + o(1)\big] \sum_{i=1}^{\boldsymbol{n}} \boldsymbol{G}(i)\boldsymbol{\mu}(i) = \big[1 + o(1)\big] \sum_{i=1}^{\boldsymbol{n}} (\boldsymbol{L}\boldsymbol{F})(i)\boldsymbol{\mu}(i).$$

To complete the proof, it remains to recall the statement of Lemma 7.1 □

Let $\overline{g}_\varepsilon : \mathbb{T} \to \mathbb{R}$ be given by

$$\overline{g}_\varepsilon = g - r(\varepsilon)\chi_{\mathcal{E}_1},$$

where $r(\varepsilon) = E_{\mu_\varepsilon}[g]/\mu_\varepsilon(\mathcal{E}_1)$ and $E_{\mu_\varepsilon}[g]$ represents the expectation of g with respect to μ_ε. Clearly, $E_{\mu_\varepsilon}[\overline{g}_\varepsilon] = 0$, and, by Assertion 9.A and (7.5), $\lim_\varepsilon r(\varepsilon) = 0$. The following proposition is the main result of this section.

Proposition 9.1. *Let $f_\varepsilon : [\mathfrak{w}_1^-, 1 + \mathfrak{w}_1^-] \to \mathbb{R}$ be the function given by*

$$f_\varepsilon(x) = \boldsymbol{F}(1) + a(\varepsilon) \int_{\mathfrak{w}_1^-}^{x} e^{S(y)/\varepsilon}\, dy + \frac{1}{\varepsilon} e^{-H/\varepsilon} \int_{\mathfrak{w}_1^-}^{x} e^{S(y)/\varepsilon} \int_{\mathfrak{w}_1^-}^{y} \overline{g}_\varepsilon(z) e^{-S(z)/\varepsilon}\, dz\, dy, \tag{9.2}$$

where

$$a(\varepsilon) = \frac{1}{e^{B/\varepsilon} - 1} \frac{1}{\varepsilon} e^{-H/\varepsilon} \int_{\mathfrak{w}_1^-}^{1+\mathfrak{w}_1^-} \overline{g}_\varepsilon(y) e^{-S(y)/\varepsilon}\, dy.$$

Then, f_ε is 1-periodic, and solves the elliptic problem $\widehat{L}_\varepsilon f_\varepsilon = \overline{g}_\varepsilon$ in $\mathbb{T}$. Moreover, there exists a finite constant C_0 such that

$$\sup_{0<\varepsilon<1} \sup_{\theta \in \mathbb{T}} \big| f_\varepsilon(\theta) \big| \le C_0 \quad \text{and} \quad \lim_{\varepsilon \to 0} \sup_{\theta \in \mathcal{E}} \big| f_\varepsilon(\theta) - f(\theta) \big| = 0,$$

where $f : \mathbb{T} \to \mathbb{R}$ is given by $f = \sum_{1 \le i \le \boldsymbol{n}} \boldsymbol{F}(i)\chi_{\mathcal{E}_i}$.

The proof of this proposition is divided in several steps. In the next lemma, we show that the function f_ε is 1-periodic and solves the Poisson equation.

Lemma 9.2. *Let $g : \mathbb{T} \to \mathbb{R}$ be a bounded function which has mean zero with respect to μ_ε, and let $f_\varepsilon : [0, 1] \to \mathbb{R}$ be given by*

$$f_\varepsilon(x) = A + a(\varepsilon) \int_0^x e^{S(y)/\varepsilon}\, dy + \frac{1}{\varepsilon} \int_0^x e^{S(y)/\varepsilon} \int_0^y g(z) e^{-S(z)/\varepsilon}\, dz\, dy,$$

where $A \in \mathbb{R}$ and

$$a(\varepsilon) = \frac{1}{e^{B/\varepsilon} - 1} \frac{1}{\varepsilon} \int_0^1 g(y) e^{-S(y)/\varepsilon}\, dy.$$

Then, f_ε solves the elliptic problem $L_\varepsilon f = g$ in $\mathbb{T}$.

Proof. We have to show that $(L_\varepsilon f_\varepsilon)(x) = g(x)$ for all $x \in (0, 1)$ and that $f_\varepsilon'(1) = f_\varepsilon'(0)$, $f_\varepsilon(1) = f_\varepsilon(0)$. The first two properties are straightforward. The third one is proved in Assertion 9.B below. □

Assertion 9.B. *We claim that $f_\varepsilon(1) = f_\varepsilon(0)$.*

Proof. In view of its definition, $\varepsilon[f_\varepsilon(1) - f_\varepsilon(0)]$ is equal to

$$\varepsilon a(\varepsilon) \int_0^1 e^{S(y)/\varepsilon}\, dy + \int_0^1 e^{S(y)/\varepsilon} \int_0^y g(z) e^{-S(z)/\varepsilon}\, dz\, dy.$$

Change the order of integration in the second term, and recall the definition of $a(\varepsilon)$. Since $B = S(0) - S(1)$, this expression is equal to $[e^{B/\varepsilon} - 1]^{-1} e^{-S(1)/\varepsilon}$ times

$$e^{S(1)/\varepsilon} \int_0^1 g(z) e^{-S(z)/\varepsilon}\, dz \int_0^1 e^{S(y)/\varepsilon}\, dy$$
$$+ \big[e^{S(0)/\varepsilon} - e^{S(1)/\varepsilon}\big] \int_0^1 g(z) e^{-S(z)/\varepsilon} \int_z^1 e^{S(y)/\varepsilon}\, dy\, dz.$$

This difference is equal to

$$e^{S(1)/\varepsilon} \int_0^1 g(z) e^{-S(z)/\varepsilon} \int_0^z e^{S(y)/\varepsilon}\, dy\, dz$$
$$+ e^{S(0)/\varepsilon} \int_0^1 g(z) e^{-S(z)/\varepsilon} \int_z^1 e^{S(y)/\varepsilon}\, dy\, dz. \tag{9.3}$$

Rewrite the second integral as

$$\int_0^1 g(z) e^{-S(z)/\varepsilon} \int_z^{1+z} e^{S(y)/\varepsilon}\, dy\, dz - \int_0^1 g(z) e^{-S(z)/\varepsilon} \int_1^{1+z} e^{S(y)/\varepsilon}\, dy\, dz.$$

Note that in the first integral the density π_ε appears. Since g has mean zero with respect to μ_ε, the first term vanishes. In the second integral, change variables $y = y' + 1$ and recall that $S(y' + 1) = S(y') - B$ to obtain that the second integral of (9.3) is equal to

$$-e^{[S(0)-B]/\varepsilon} \int_0^1 g(z) e^{-S(z)/\varepsilon} \int_0^z e^{S(y)/\varepsilon}\, dy\, dz.$$

Since $S(0) - B = S(1)$, the terms in (9.3) cancel, which completes the proof of the assertion. □

Proof of Proposition 9.1. We proved in Lemma 9.2 that f_ε is 1-periodic and solves the elliptic equation $\widehat{L}_\varepsilon f_\varepsilon = \overline{g}_\varepsilon$ in $\mathbb{T}$. It remains to show that f_ε is uniformly bounded and converges uniformly to f on the set $\mathcal{E}$. We examine separately the second and third terms on the right-hand side of (9.2).

We claim that the second term vanishes as $\varepsilon \to 0$, uniformly in $x \in [\mathfrak{w}_1^-, 1 + \mathfrak{w}_1^-]$. On the one hand, since $S(x) \le S(\mathfrak{w}_1^-)$ for all $x \ge \mathfrak{w}_1^-$ [because $\mathfrak{w}_1^-$ is the left endpoint of a valley],

$$\frac{1}{\sqrt{\varepsilon}} \int_{\mathfrak{w}_1^-}^{1+\mathfrak{w}_1^-} e^{[S(y)-S(\mathfrak{w}_1^-)]/\varepsilon}\, dy \le C_0 \tag{9.4}$$

for some finite constant C_0 independent of ε. On the other hand, since $|\overline{g}_\varepsilon(x)| \le C_0$, $e^{B/\varepsilon} - 1 \ge (1 - e^{-B}) e^{B/\varepsilon}$ for sufficiently small $\varepsilon > 0$, and $S(1 + \mathfrak{w}_1^-) = S(\mathfrak{w}_1^-) - B$,

$$\frac{1}{e^{B/\varepsilon} - 1} e^{[S(\mathfrak{w}_1^-)-H]/\varepsilon} \frac{1}{\sqrt{\varepsilon}} \left| \int_{\mathfrak{w}_1^-}^{1+\mathfrak{w}_1^-} \overline{g}_\varepsilon(y) e^{-S(y)/\varepsilon}\, dy \right|$$
$$\le C_0 e^{[S(1+\mathfrak{w}_1^-)-H]/\varepsilon} \frac{1}{\sqrt{\varepsilon}} \int_{\mathfrak{w}_1^-}^{1+\mathfrak{w}_1^-} e^{-S(y)/\varepsilon}\, dy. \tag{9.5}$$

Since $\mathfrak{w}_1^-$ is the left endpoint of a valley whose depth is H, $S(1+\mathfrak{w}_1^-) - H = S(1+\mathfrak{m}_{1,1})$. By (9.1), there exists $\eta > 0$ such that $S(y) \ge S(1+\mathfrak{m}_{1,1}) + \eta$ for all $\mathfrak{w}_1^- \le y \le 1+\mathfrak{w}_1^-$. Multiplying (9.4) and (9.5) yields that the second term in the formula of f_ε vanishes as $\varepsilon \to 0$, uniformly for $x \in [\mathfrak{w}_1^-, 1+\mathfrak{w}_1^-]$.

We turn to the third term on the right hand side of (9.2). Exchange the order of the integrals to write it as

$$\frac{1}{\varepsilon} e^{-H/\varepsilon} \int_{\mathfrak{w}_1^-}^{x} \overline{g}_\varepsilon(z) \int_z^x e^{[S(y)-S(z)]/\varepsilon}\, dy\, dz. \tag{9.6}$$

Since $\overline{g}_\varepsilon$ is uniformly bounded, the absolute value of this expression is bounded by

$$\frac{C_0}{\varepsilon} e^{-H/\varepsilon} \int_{\mathfrak{w}_1^-}^{1+\mathfrak{w}_1^-} \int_z^{1+z} e^{[S(y)-S(z)]/\varepsilon}\, dy\, dz = \frac{C_0}{\varepsilon} e^{-H/\varepsilon} c(\varepsilon),$$

where $c(\varepsilon)$ is the normalizing constant introduced below (2.2). By (5.10), this expression is uniformly bounded in ε. This proves one assertion of the proposition. It also proves that we may replace $\overline{g}_\varepsilon$ in (9.6) by g because, by Assertion 9.A, $r(\varepsilon)$ converges to 0 as $\varepsilon \to 0$.

It remains to prove the uniform convergence in $\mathcal{E}$ of the sequence f_ε with $\overline{g}_\varepsilon$ replaced by g. As the integral in (9.6) is carried over pairs (y, z) such that $z \le y$ the maximum value of the difference $S(y) - S(z)$ is H and it is attained only when y, z belong to the same landscape and $V(y) = H$, $V(z) = 0$, that is, when y is an endpoint of a valley, and z is a global minima of a valley in the same landscape. Hence, the dominant terms of the integral are the ones in which y belongs to a neighborhood of an endpoint of a valley and z to a neighborhood of a global minima of a valley.

Recall from Section 7.2 that the valleys $\mathcal{W}_j$ which belong to the same landscape are represented as $\mathcal{W}_{a,1}, \dots, \mathcal{W}_{a,n_a}$, $1 \le a \le \boldsymbol{p}$. To prove uniform convergence in $\mathcal{E}$, fix a point $x \in \mathcal{E}_j = \mathcal{E}_{a,\ell}$. By the observation of the previous paragraph, the contribution to the integral of the points z which do not belong to a neighborhood of point $\mathfrak{m}_{b,k,i}$ [that is a global minimum of V in the valley $\mathcal{W}_{b,k}$] is negligible.

Fix $b < a$, $1 \le k \le n_b$. The contribution to integral when z belongs to the neighborhoods of the local minima of $\mathcal{W}_{b,k}$ is given by $\boldsymbol{\pi}(b,k) G_1(\mathfrak{m}_{b,k,1})$, where $\boldsymbol{\pi}$ has been introduced in (7.1). For $b = a$, $1 \le k \le \ell \le n_a$, the contribution to integral of the neighborhoods of the local minima of $\mathcal{W}_{a,k}$ is equal to $\boldsymbol{\pi}(a,k)[G_1(\mathfrak{m}_{a,k,1}) - G_1(\mathfrak{m}_{a,\ell,1})]$. Hence, the integral (9.6) with $\overline{g}_\varepsilon$ replaced by g is equal to

$$\begin{aligned} &\sum_{b=1}^{a-1} \sum_{k=1}^{n_b} (\boldsymbol{L}\boldsymbol{F})(b,k) \boldsymbol{\pi}(b,k) G_1(\mathfrak{m}_{b,k,1}) \\ &\quad + \sum_{k=1}^{\ell-1} (\boldsymbol{L}\boldsymbol{F})(a,k) \boldsymbol{\pi}(a,k) \big[G_1(\mathfrak{m}_{a,k,1}) - G_1(\mathfrak{m}_{a,\ell,1})\big] + R(\varepsilon), \end{aligned}$$

where $R(\varepsilon)$ is a remainder which converges to 0 as $\varepsilon \to 0$, uniformly for $x \in [\mathfrak{w}_1^-, 1+\mathfrak{w}_1^-]$. By Lemma 7.7, the previous sum is equal to $\boldsymbol{F}(a,\ell) - \boldsymbol{F}(1)$, which proves that f_ε converges to f uniformly in $\mathcal{E}$. □

Acknowledgements

The authors wish to thank the anonymous referee for her/his comments. The first author has been partially supported by FAPERJ CNE E-26/201.207/2014, by CNPq Bolsa de Produtividade em Pesquisa PQ 303538/2014-7, by ANR-15-CE40-0020-01 LSD of the French National Research Agency. I. Seo was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2018R1C1B6006896).

References

[1] J. Beltrán and C. Landim. Tunneling and metastability of continuous time Markov chains. *J. Stat. Phys.* **140** (2010) 1065–1114. MR2684500
[2] J. Beltrán and C. Landim. Tunneling and metastability of continuous time Markov chains II, the nonreversible case. *J. Stat. Phys.* **149** (2012) 598–618. MR2998592

[3] J. Beltrán and C. Landim. A martingale approach to metastability. *Probab. Theory Related Fields* **161** (2015) 267–307. MR3304753
[4] F. Bouchet and J. Reygner. Generalisation of the Eyring–Kramers transition rate formula to irreversible diffusion processes. *Ann. Henri Poincaré* **17** (2016) 3499–3532. MR3568024
[5] A. Bovier, M. Eckhoff, V. Gayrard and M. Klein. Metastability in reversible diffusion process I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)* **6** (2004) 399–424. MR2094397
[6] L. C. Evans and P. R. Tabrizian. Asymptotic for scaled Kramers–Smoluchowski equations. *SIAM J. Math. Anal.* **48** (2016) 2944–2961. MR3542005
[7] A. Faggionato and D. Gabrielli. A representation formula for large deviations rate functionals of invariant measures on the one dimensional torus. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** (2012) 212–234. MR2919204
[8] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*, 2nd edition. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **260**. Springer-Verlag, New York, 1998. MR1652127
[9] A. Gaudillière and C. Landim. A Dirichlet principle for non reversible Markov chains and some recurrence theorems. *Probab. Theory Related Fields* **158** (2014) 55–89. MR3152780
[10] P. Imkeller and I. Pavlyukevich. Metastable behaviour of small noise Lévy-driven diffusions. *ESAIM Probab. Stat.* **12** (2008) 412–437. MR2437717
[11] C. Landim. Metastability for a non-reversible dynamics: The evolution of the condensate in totally asymmetric zero range processes. *Comm. Math. Phys.* **330** (2014) 1–32. MR3215575
[12] C. Landim, M. Loulakis and M. Mourragui. Metastable Markov chains: From the convergence of the trace to the convergence of the finite-dimensional distributions. *Electron. J. Probab.* **23** (2018), 95.
[13] C. Landim, M. Mariani and I. Seo. A Dirichlet and a Thomson principle for non-selfadjoint elliptic operators, metastability in non-reversible diffusion processes. *Arch. Ration. Mech. Anal.* **231** (2019) 887–938. Available at https://doi.org/10.1007/s00205-018-1291-8.
[14] C. Landim and I. Seo. Metastability of non-reversible, mean-field Potts model with three spins. *J. Stat. Phys.* **165** (2016) 693–726. MR3568163
[15] C. Landim and I. Seo. Metastability of non-reversible random walks in a potential field, the Eyring–Kramers transition rate formula. *Comm. Pure Appl. Math.* **LXXI** (2018) 203–266. MR3745152
[16] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales: Volume 1, Foundations*. Cambridge University Press, Cambridge, 1994. MR1780932
[17] I. Seo and P. R. Tabrizian. Asymptotics for scaled Kramers–Smoluchowski equations in several dimensions with general potentials, 2018. Available at arXiv:1808.09108.
[18] M. Slowik. A note on variational representations of capacities for reversible and nonreversible Markov chains Technische Universität Berlin, 2012. Unpublished.
[19] M. Sugiura. Metastable behaviors of diffusion processes with small parameter. *J. Math. Soc. Japan* **47** (1995) 755–788. MR1348758